

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

FOURIE  
KARLSSON

88735

3/13 08

Access DB# 3/13 291

# SEARCH REQUEST FORM

Scientific and Technical Information Center

CRFE

Requester's Full Name: M.A. WALICKA Examiner #: 78201 Date: March 11, '03  
Art Unit: 1652 Phone Number 305-7370 Serial Number: 101050,200  
Mail Box and Bldg/Room Location: 10D06 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Aggrecanase-1 and -2 peptides and methods  
Inventors (please provide full names): FOURIE A. et al.

Earliest Priority Filing Date: Jan. 18, 2002 J 16 2002

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Point of Contact:  
Toby Port  
Technical Info. Specialist  
CM1 EA04  
703-300-3534

Please search:

SEQ ID NO: 3

4

5 PA

6

7

polypeptides.

AA Thank you in advance.

3-11

4-11

5-10

6-10

7-13

Articles

RECEIVED  
MAR 11 2003  
STIC

## STAFF USE ONLY

Searcher:	Type of Search	Vendors and cost where applicable
Searcher Phone #:	NA Sequence (#)	STN
Searcher Location:	AA Sequence (#) <u>5</u>	Dialog
Date Searcher Picked Up: <u>3/12</u>	Structure (#)	Questel/Orbit
Date Completed: <u>3/14</u>	Bibliographic	Dr. Link
Searcher Prep & Review Time: <u>12</u>	Litigation	Lexis/Nexis
Clerical Prep Time:	Fulltext	Sequence Systems <u>es</u>
Online Time: <u>10</u>	Patent Family	WWW/Internet
	Other	Other (specify)

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2003, 22:43:04 ; Search time 36.1818 seconds  
(without alignments)  
36.828 Million cell updates/sec

Title: PCT-US03-01327-5  
Perfect score: 44  
Sequence: 1 EKARVLAEMA 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_101002:\*

1:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	10	23	ABR80017
2	44	100.0	11	23	ABR80016
3	44	100.0	11	23	ABR6233
4	38	86.4	219	23	AA083139
5	38	86.4	396	22	AA067257
6	38	86.4	396	22	AA094593
7	36	81.8	298	21	AA074531
8	36	81.8	298	21	AA074532
9	36	81.8	298	21	AA074533
10	35	79.5	8	22	ABP12631

*Cathepsin S is a cysteine protease*

11	35	79.5	8	22	ABP20832
12	35	79.5	9	22	ABP12751
13	35	79.5	9	22	ABP17961
14	35	79.5	10	23	AAE19118
15	35	79.5	10	23	AAU75795
16	35	79.5	11	22	ABP17190
17	35	79.5	12	18	AAW36881
18	35	79.5	12	22	ABP17219
19	35	79.5	12	22	ABP17220
20	35	79.5	12	22	ABP25633
21	35	79.5	15	16	AAE68705
22	35	79.5	15	18	AAW25875
23	35	79.5	15	19	AAW67391
24	35	79.5	15	20	AAV39674
25	35	79.5	15	20	AAW98894
26	35	79.5	15	22	ABP24909
27	35	79.5	15	22	ABP24915
28	35	79.5	24	11	AAW07981
29	35	79.5	25	19	AAW82529
30	35	79.5	30	23	AAU84422
31	35	79.5	32	16	AAE68676
32	35	79.5	32	18	AAW25846
33	35	79.5	32	19	AAW67362
34	35	79.5	32	20	AAV39767
35	35	79.5	32	20	AAW99970
36	35	79.5	34	16	AAE68706
37	35	79.5	34	16	AAE68707
38	35	79.5	34	16	AAE77126
39	35	79.5	34	18	AAW25877
40	35	79.5	34	18	AAW25876
41	35	79.5	34	19	AAW67392
42	35	79.5	34	19	AAW67393
43	35	79.5	34	20	AAV39722
44	35	79.5	34	20	AAV39723
45	35	79.5	34	20	AAW98895

## ALIGNMENTS

RESULT 1		
ABBB0017	ID	ABBB0017 standard; peptide; 10 AA.
XX	AC	ABBB0017;
XX	DT	17-JUL-2002 (first entry)
XX	DE	Cathepsin S inhibition assay substrate peptide.
KW	KW	Cathepsin S; antiallergic; antisthmatic; dermatological; pyrazole;
KW	KW	atopic allergy; hay fever; asthma; atopic dermatitis; food allergy;
XX	XX	allergy; dust; pollen; mold; pet dander; pet hair.
OS	OS	Unidentified.
XX	FT	Key
XX	FT	Modified-site
XX	XX	Location/Qualifiers
XX	XX	10
XX	XX	/note="C-terminal amide"
XX	XX	WO200220011-A2.
XX	PD	14-MAR-2002.
XX	XX	
XX	FE	05-SEP-2001; 2001MO-US27429.
XX	XX	
XX	PR	06-SEP-2000; 2000US-230407P.
XX	PR	10-AUG-2001; 2001US-0927324.
XX	XX	
XX	PA	(ORTH ) ORTHO-MCNELL PHARM INC.
XX	XX	
PI	PI	Butler CR, Cai H, Edwards JP, Grice CA, Gu Y, Gustin DJ,
PI	PI	Karlsson L, Khutuya H, Medina SP, Pio BA, Sehon CA, Sun S;

HIV A03 motif gag  
HIV A02 super moti  
HIV B56 super moti  
Protease peptide.  
Human immunodefici  
HIV B27 super moti  
HIV B27 super cleav  
HIV B27 super moti  
HIV B27 super moti  
HIV B27 super moti  
HIV B27 super moti  
T cell epitope, p2  
HIV T-helper deter  
HIV-1 peptide epit  
HIV1 gag protein e  
HIV-1 vaccine synt  
HIV DR 3a motif ga  
Gag protein T-Cell  
HIV-1 p24 epitope  
HIV GAG segment 24  
P24M/V3 consensus  
Chimeric T/B cell  
HIV-1 peptide epit  
HIV1 chimeric pept  
P24M/CURB-56 chime  
CURB-56/P24M chime  
P24M/V3 consensus  
Chimeric T/B cell  
Chimeric T/B cell  
HIV-1 peptide epit  
HIV-1 peptide epit  
HIV1 chimeric pept  
HIV1 chimeric pept  
HIV-1 vaccine synt

**THIS PAGE BLANK (USP 12)**

pct-us03-01327-5.regular.rag

```

PS      Example 37; Page 92; 115pp; English.
XX
CC      The invention relates to a method of treating allergic conditions
CC      comprising administration of a composition comprising a substituted
CC      pyrazole (I) or a salt, amide or ester. The method is used for treating
CC      allergic conditions, e.g. hay fever, asthma, atopic dermatitis and food
CC      allergies. The present sequence represents human cathepsin S
CC      substrate peptide used in an assay to demonstrate the method of the
CC      invention.
XX
XX      Sequence 11 AA:
SQ
Query Match 100.0%; Score 44; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY.      1 EKARYLAEEA 10
          |||||
DB        1 EKARYLAEEA 10

```

RESULT 3	
ABB76233	
ID ABB76233	standard: Peptide: 11 AA..

AC	ABB76233;
XX	
DT	09-AUG-2002 (first entry)
DT	

XX Cathepsin S; inhibitor; hay fever; asthma; food allergy;  
KW atopic dermatitis; antiallergic; antisthmatic; dermatological;  
KW

XX

XX  
XX

XXXXX

FT	Modified-site	1	Reference
FT	Modified-site	1	Reference

```
FT      .Modified-site      11      "neighbor" justno-emptdo"
FT
```

/note= "Dabcyyl lysine-amide"

PN W0200220002-A2

PD 14-MAR-2002.

05-SEP-2001; 2001WO-US27441

PR 06-SEP-2000; 2000US-230407P.

PA (ORTH ) ORTHO-MCNEIL PHARM INC

PI Gu Y, Karlsson L, Sun S, Thurmond RL,

DR. WPI; 2002-443877/47

PT Use of cathepsin S inhibitor for treating an allergic condition, e.g. hay fever

[illegible]

XX  
XX  
3

CC cathepsin S inhibition assay. The  $K_m$  for the substrate is around  
CC 5 mM but the presence of substrate inhibition makes kinetic

analysis difficult. With 20  $\mu\text{M}$  substrate, the assay rate is linear over the range of 1–8  $\mu\text{M}$  substrate. At 100  $\mu\text{M}$  reaction

CC Using 2 ng/well of cathepsin S, the production of product is

CC loss of substrate. Primary assays are run by quenching the CC reaction after 20 min and then measuring the fluorescence. For

CC rate is calculated from the slope of the increase and the percent

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:40:08 ; Search time 26.5455 Seconds  
(without alignments)  
50.197 Million cell updates/sec

Title: US-10-050-200-5  
Perfect score: 44  
Sequence: 1 EKARYLAEMA 10

Scoring table: BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq-101002:\*

1:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	100.0	10	23	ABB80017	Cathepsin S inhibi
2	44	100.0	11	23	AA080096	Human cathepsin S
3	44	100.0	11	23	ABB76233	Substrate for cath
4	38	86.4	219	23	AA083139	Novel secreted pro
5	38	86.4	396	22	AA67257	Amino acid sequenc
6	38	86.4	396	22	AA094593	Human protein sequ
7	36	81.8	298	21	AA174531	Neisseria gonorrhoe
8	36	81.8	298	21	AA174532	Neisseria meningit
9	36	81.8	298	21	AA174533	Neisseria meningit
10	35	79.5	22	22	ABP12631	HIV A02 super moti

11	35	79.5	8	22	ABP20832	HIV A03 motif gag
12	35	79.5	9	22	ABP12751	HIV A04 super moti
13	35	79.5	9	22	ABP17961	HIV B58 super moti
14	35	79.5	10	23	AAE19118	Protease peptide
15	35	79.5	10	23	AAU75795	Human Immunodefici
16	35	79.5	11	22	ABP17190	HIV B27 super moti
17	35	79.5	12	18	AAW36881	HIV protease cleav
18	35	79.5	12	22	ABP17219	HIV B27 super moti
19	35	79.5	12	22	ABP17220	HIV B27 super moti
20	35	79.5	14	13	AAE25533	HIV1, gag, 357-370
21	35	79.5	15	16	AAE68705	T cell epitope, P2
22	35	79.5	15	15	AAW25875	HIV T-helper deter
23	35	79.5	15	19	AAW67391	HIV-1 peptide epit
24	35	79.5	15	20	AAW39674	HIV1 gag protein e
25	35	79.5	15	20	AAW98894	HIV-1 vaccine synt
26	35	79.5	15	22	ABP24809	HIV DR 3a motif ga
27	35	79.5	15	22	ABP24815	HIV DR 3a motif ga
28	35	79.5	24	11	AAW07981	Gag protein T-Cell
29	35	79.5	25	19	AAW82529	HIV-1 p24 epitope
30	35	79.5	30	23	AAU84422	HIV GAG segment 24
31	35	79.5	32	16	AAE68676	P24W/V3 consensus
32	35	79.5	32	18	AAW25846	Chimeric T/B cell
33	35	79.5	32	19	AAW67362	HIV-1 peptide epit
34	35	79.5	32	20	AAW39767	HIV1 chimeric pept
35	35	79.5	32	20	AAW99970	HIV-1 vaccine synt
36	35	79.5	34	16	AAE68706	P24W/CTRB-56 chime
37	35	79.5	34	16	AAE68707	CTRB-56/P24M chime
38	35	79.5	34	16	AAE77726	P24W/V3 consensus
39	35	79.5	34	18	AAW25877	Chimeric T/B cell
40	35	79.5	34	18	AAW25876	Chimeric T/B cell
41	35	79.5	34	19	AAW67392	HIV-1 peptide epit
42	35	79.5	34	19	AAW67393	HIV-1 peptide epit
43	35	79.5	34	20	AAW39722	HIV1 chimeric pept
44	35	79.5	34	20	AAW39723	HIV1 chimeric pept
45	35	79.5	34	20	AAW98895	HIV-1 vaccine synt

#### ALIGNMENTS

RESULT 1						
ID	ABB80017	standard; peptide; 10 AA.				
XX	ABB80017:					
DT	17-JUL-2002	(first entry)				
DE	Cathepsin S inhibition assay substrate peptide.					
XX						
KW	Cathepsin S; antiallergic; antisthmatic; dermatological; pyrazole;					
KW	atopic allergy; hay fever; asthma; atopic dermatitis; food allergy;					
KW	allergy; dust; pollen; mold; pet dander; pet hair.					
XX						
OS	Unidentified.					
XX						
FH	Key	Location/Qualifiers				
FT	Modified-site	10				
XX		/note="C-terminal amide"				
PN	WO200220011-A2.					
PD	14-MAR-2002.					
XX						
PF	05-SEP-2001; 2001WO-US27429.					
XX						
PR	06-SEP-2000; 2000US-230407P.					
XX						
PR	10-AUG-2001; 2001US-092732A.					
XX						
PA	(ORTH ) ORTHO-MCNETIL PHARM INC.					
XX						
PI	Butler CR, Cai H, Edwards JP, Grice CA, Gu Y, Gustin DJ;					
PI	Karlsson L, Khatuya H, Medina SP, Pio BA, Sehon CA, Sun S;					

PI Tays KL, Thurmond RL, Wel J;  
 XX  
 DR WPI: 2002-393814/42.  
 XX  
 PT Treatment of allergic conditions e.g. atopic dermatitis or asthma  
 PT comprises administration of pyrazole derivative that inhibits cathepsin  
 S  
 XX  
 PS Example 54; Page 134; 165pp; English.  
 XX  
 CC The invention relates to the treatment of allergic conditions, comprising  
 CC the administration of a pyrazole derivative. The activity of derivatives  
 CC of the invention may be described as, antiallergic, antiasthmatic and  
 CC dermatological. They act as cathepsin S inhibitors and may be used for  
 CC the treatment of allergic conditions including atopic allergies, e.g. hay  
 CC fever, asthma, atopic dermatitis, food allergies and allergies to dust,  
 CC pollen, mold, pet dander or pet hair. The current sequence represents a  
 CC cathepsin S inhibition assay substrate peptide.  
 XX  
 SQ Sequence 10 AA:  
 Query Match 100.0%; Score 44; DB 23; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.031;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 EKARVLAEEA 10  
 DB 1 EKARVLAEEA 10  
 .  
 RESULT 2  
 AAU80096  
 ID AAU80096 standard; peptide: 11 AA.  
 XX  
 AC AAU80096;  
 XX  
 DT 24-SEP-2002 (first entry)  
 XX  
 DE Human cathepsin S substrate peptide.  
 XX  
 DE Cathepsin S; allergy: pyrazole; hay fever; asthma; atopic dermatitis;  
 KM food allergy.  
 KM  
 OS Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "Aedens-Glu"  
 FT Modified-site 10 /note= "Dabcy1"  
 FT Modified-site 11 /label= C-terminal amide  
 FT  
 XX  
 PA WO200220013-A2.  
 PN  
 PD 14-MAR-2002.  
 PD  
 PF 05-SEP-2001; 2001WO-US27480.  
 PF  
 PR 06-SEP-2000; 2000US-230407P.  
 PR 10-AUG-2001; 2001US-0927188.  
 XX  
 PA (ORTH ) ORTHO-MCNEIL PHARM INC.  
 XX  
 PI Cat H, Edwards JP, Gu Y, Karlsson L, Medina SP, Pio BA, Sun S;  
 PI Thurmond RL, Wel J;  
 XX  
 DR WPI: 2002-527344/56.  
 XX  
 PT Treating allergic conditions, e.g. hay fever, asthma, atopic dermatitis  
 PT and food allergies, comprises administration of a substituted pyrazole  
 PT  
 XX

PS Example 37; Page 92; 115pp; English.  
 XX  
 CC The invention relates to a method of treating allergic conditions  
 CC comprising administration of a composition comprising a substituted  
 CC pyrazole (1) or a salt, amide or ester. The method is used for treating  
 CC allergic conditions, e.g. hay fever, asthma, atopic dermatitis and food  
 CC allergies. The present sequence represents human cathepsin S  
 CC substrate peptide used in an assay to demonstrate the method of the  
 CC invention.  
 XX  
 SQ Sequence 11 AA:  
 Query Match 100.0%; Score 44; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 EKARVLAEEA 10  
 DB 1 EKARVLAEEA 10  
 .  
 RESULT 3  
 ABB76233  
 ID ABB76233 standard; Peptide: 11 AA.  
 XX  
 AC ABB76233;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Substrate for cathepsin S inhibition assay.  
 XX  
 DE Cathepsin S; inhibitor: allergy; hay fever; asthma; food allergy;  
 KM atopic dermatitis; antiallergic; antiasthmatic; dermatological;  
 KM antiinflammatory; immunosuppressive.  
 KM  
 OS Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "Aedens-glutamic acid"  
 FT Modified-site 11 /note= "Dabcy1 lysine-amide"  
 FT  
 XX  
 PA WO200220002-A2.  
 PN  
 PD 14-MAR-2002.  
 PD  
 PF 05-SEP-2001; 2001WO-US27441.  
 PF  
 PR 06-SEP-2000; 2000US-230407P.  
 PR  
 PA (ORTH ) ORTHO-MCNEIL PHARM INC.  
 XX  
 PI Gu Y, Karlsson L, Sun S, Thurmond RL;  
 PI WPI: 2002-443877/47.  
 DR  
 DR WPI: 2002-443877/47.  
 DR  
 PT Use of cathepsin S inhibitor for treating an allergic condition, e.g.  
 PT hay fever, asthma, atopic dermatitis or a food allergy  
 PT  
 PS Example 1; Page 29; 53pp; English.  
 XX  
 CC The present sequence is a peptide used as substrate for a  
 CC cathepsin S inhibition assay. The Km for the substrate is around  
 CC 5 uM but the presence of substrate inhibition makes kinetic  
 CC analysis difficult. With 20 uM substrate, the assay rate is  
 CC linear over the range of 1-8 ng cathepsin S in 100 ul reaction.  
 CC using 2 ng/well of cathepsin S, the production of product is  
 CC linear and yields about 7-fold signal after 20 min with only 20%  
 CC loss of substrate. Primary assays are run by quenching the  
 CC reaction after 20 min and then measuring the fluorescence. For  
 CC other assays, measurements are taken every min for 20 min. The  
 CC rate is calculated from the slope of the increase and the percent

CC Inhibition is calculated from this. The present invention relates  
 CC to the use of cathepsin S inhibitors to treat an allergic  
 CC condition, especially an atopic IgE-mediated allergic condition  
 CC such as hay fever, asthma, atopic dermatitis or a food allergy  
 CC (claimed). The allergens may include dust, pollen, mould, and pet  
 CC dander or pet hair.

XX Sequence 11 AA;

Query Match 100.0%; Score 44; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKARVLAEEA 10  
 |||:|||||  
 Db 1 EKARVLAEEA 10

#### RESULT 4

ID AAU83139 standard; Protein; 219 AA.

AC AAU83139;

DT 08-MAY-2002 (first entry)

DE Novel secreted protein Z838027G3P.

KW Protein secretion; mammalian secreted polypeptide; MSP.

XX Homo sapiens.

PN W0200202621-A2.

PD 10-JAN-2002.

PF 28-JUN-2001; 2001WO-US20638.

PR 30-JUN-2000; 2000US-215446P.

XX (ZYMO) ZYMOGENETICS INC.

PI Sheppard PO, Presnell SR;

DR WPI; 2002-147999/19.

DR N-PSDB; ABK33054.

PT Novel isolated mammalian secreted polypeptide useful in therapeutic and  
 PT diagnostic methods, to direct secretion of other proteins of interest  
 PT from host cell, as educational tools, and as laboratory practicum kits

XX Claim 12; Page 194-195; 397pp; English.

CC The invention describes an isolated mammalian secreted polypeptide (MSP)  
 CC (1). (1) is useful to direct the secretion of other proteins of interest  
 CC from a host cell, to monitor secretion of proteins, to degenerate  
 CC sequences comprising all nucleotide sequences encoding a particular  
 CC polypeptide, to screen for cell metabolism effecting receptors, for  
 CC identifying new target receptors and drug design, for identifying, for  
 CC protein purification, for determining the weight of expressed MSP  
 CC polypeptides as a ratio to total protein expressed, for identifying  
 CC peptide cleavage sites, for coupling amino and carboxy terminal tags, for  
 CC amino acid sequence analysis, for monitoring biological activities of the  
 CC protein in vitro and in vivo, and to teach analytical skills and as  
 CC reagents for the study of cells, receptors, and other binding molecules.  
 CC The polynucleotide is useful for radiation hybrid mapping, and somatic  
 CC cell genetic technique developed for constructing high-resolution.  
 CC contiguous maps of mammalian chromosomes. Reagents disclosed in the  
 CC invention may be used to detect metabolic abnormalities characterised by  
 CC over or under production of the protein. This is the amino acid sequence  
 CC of a mammalian secreted polypeptide, described in the method of the  
 CC invention.

XX Sequence 219 AA;

Query Match 86.4%; Score 38; DB 23; Length 219;  
 Best Local Similarity 80.0%; Pred. No. 11;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKARVLAEEA 10  
 |||:|||||  
 Db 52 EKSRVLAEEA 61

#### RESULT 5

ID AAG67257 standard; Protein; 396 AA.

AC AAG67257;

DT 13-NOV-2001 (first entry)

DE Amino acid sequence of a human liver-associated gene.

KW Liver-associated gene; hepatic cancer; liver cancer; marker.

XX Homo sapiens.

PN W0200109318-A1.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-JP05064.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 18-OCT-1999; 99US-0159590.

PR 11-JAN-2000; 2000JP-0118776.

PR 17-FEB-2000; 2000US-0183322.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Aburatani H;

PI Kodama T, Hippo Y, Taniguchi H;

XX WPI; 2001-541221/60.

DR N-PSDB; AAH77810.

XX New genes and their proteins found in hepatic cancer, vectors,

PT antibodies, and screening method for compounds with binding activity,

PT useful as diagnostics and targets for treatment agents -

XX Claim 1; Page 79-80; 99pp; Japanese.

CC The present sequence is encoded by a liver-associated gene. The  
 CC expression level of this gene changes in liver (hepatic) cancer.  
 CC Liver-associated gene can be used as markers in blood tests for  
 CC screening for the early stages of the liver cancer. The encoded  
 CC proteins and peptides can also be used as targets for screening  
 CC for compounds to treat the disease. They can also be used for  
 CC predicting the progress of the disease.

QY 1 EKARVLAEEA 10  
 |||:|||||  
 Db 229 EKSRVLAEEA 238

```

RESULT 6
AAB94593
ID AAB94593 standard; Protein; 396 AA.
AC AAB94593;
XX
XX
XX 26-JUN-2001 (first entry)
XX
XX
XX Human protein sequence SEQ ID NO:15405.
DE
XX
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EPI074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1998; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isegai T, Nishikawa T, Hayashi K, Salto K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX PT primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX PS Claim 8; SEQ ID 15405; 2537pp + CD ROM; English.
XX
XX CC The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides, where the
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13632 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX SQ Sequence 396 AA:
XX
XX Query Match 86.4%; Score 38; DB 22; Length 396;
XX Best Local Similarity 80.0%; Pred. No. 19;
XX Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 7
AA774531
ID AA774531 standard; Protein; 298 AA.
AC AA774531;
XX
XX
XX 21-MAR-2000 (first entry)
XX
XX
XX Neisseria gonorrhoeae ORF 138 protein sequence SEQ ID NO:536.
DE
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX antibacterial; gene therapy.
XX
XX Neisseria gonorrhoeae.
XX
XX WO957280-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US09346.
XX
XX 01-MAY-1998; 98US-0083758.
XX
XX 31-JUL-1998; 98US-0094869.
XX
XX 02-SEP-1998; 98US-0098994.
XX
XX 02-SEP-1998; 98US-0099062.
XX
XX 09-OCT-1998; 98US-0103749.
XX
XX 09-OCT-1998; 98US-0103794.
XX
XX 09-OCT-1998; 98US-0103796.
XX
XX 25-FEB-1999; 99US-0121528.
XX
XX (CHIR ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
XX Petersen J, Pizsa M, Rappold R, Ratti G, Scalato E, Scarselli M;
XX Tettelin H, Venter JC;
XX WPI; 2000-062150/05.
XX N-PSDB; AA253293.
XX
XX DR Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics -
XX
XX PS Claim 2; Page 390; 1453pp; English.
XX
XX AA253015 to AA254536, AA254577 to AA254615, and AA774253 to AA775941
XX represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
XX and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of
XX the invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the
XX manufacture of medicaments for treating or preventing infection due to
XX Neisserial bacteria (e.g. meningitis and septicemia), to detect the
XX presence of Neisseria bacteria, or to raise antibodies. They may also
XX be used to screen for agonists or antagonists, which may themselves
XX have use as antibacterial agents. The polynucleotides of the invention
XX may also be used in gene therapy protocols.
XX
XX SQ Sequence 298 AA:
XX
XX Query Match 81.8%; Score 36; DB 21; Length 298;
XX Best Local Similarity 77.8%; Pred. No. 36;
XX Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX OY 1 EKARVLAEA 9
XX :|||:||||
XX DB 11 DKARITLAEA 19

```



```

RESULT 8
AAV74532
ID AAV74532 standard; Protein; 298 AA.
XX
AC AAV74532;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 138 protein sequence SEQ ID NO:538.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
PN WO957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR MPI: 2000-062150/05.
DR N-PSDB; AA253294.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 2; Page 390-391; 1453pp; English.
XX
CC AA25015 to AA254536, AA254577 to AA254615, and AAV74253 to AAV75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 298 AA;
XX
Query Match 81.8%; Score 36; DB 21; Length 298;
Best Local Similarity 77.8%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 EKARVIAEA 9
:|||||
DB 11 DKARIIAEA 19

```

```

ID AAV74533 standard; Protein; 298 AA.
XX
AC AAV74533;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 138 protein sequence SEQ ID NO:540.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
PN WO957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR MPI: 2000-062150/05.
DR N-PSDB; AA253295.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 2; Page 392; 1453pp; English.
XX
CC AA25015 to AA254536, AA254577 to AA254615, and AAV74253 to AAV75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 298 AA;
XX
Query Match 81.8%; Score 36; DB 21; Length 298;
Best Local Similarity 77.8%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 EKARVIAEA 9
:|||||
DB 11 DKARIIAEA 19

```

RESULT 9  
AAV74533

RESULT 10  
ABP12631  
ID ABP12631 standard; Peptide; 8 AA.  
XX

```

AC  ABP12631;
XX
XX  15-JUL-2002 (first entry)
XX
DE  HIV A02 super motif gag peptide #102.
XX
XX  HIV; HIV-1; human immunodeficiency virus: env; pol; gag; nef; vpr;
XX  vpu; vif; tat; cytoxic T lymphocyte; CTL; immune response; epitope;
XX  antigen; vaccine; HIV infection; immunisation; virucide.
XX
OS  Human immunodeficiency virus type 1.
XX
XX  WO200124810-A1.
XX
XX  12-APR-2001.
XX
XX  05-OCT-2000; 2000WO-US27766.
XX
XX  05-OCT-1999; 99US-0412863.
XX
XX  (EPI-M-) EPIMUNE INC.
XX
XX  Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX  Baker DM, Cells E, Kubo RT, Grey HM;
XX  WPI: 2001-354887/37.
XX
XX  Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
XX  peptide groups, useful for vaccinating against HIV-1.
XX
XX  Claim 32; Page 127; 448pp; English.
XX
XX  The present invention describes a composition (I) comprising a prepared
XX  human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
XX  sequence selected from 51 defined amino acid sequences (AB125347 to
XX  ABP25397). (I) has virucide activity and can be used in vaccines. (I)
XX  may be used for immunising subjects against HIV-1 infections. The use of
XX  group-based vaccines has several advantages over traditional vaccines,
XX  particularly when compared to the use of whole antigens in vaccine
XX  compositions. There is evidence that the immune response to whole
XX  antigens is directed largely toward variable regions of the antigen,
XX  allowing for immune escape due to mutations. The groups for inclusion in
XX  an group-based vaccine may be selected from conserved regions of viral or
XX  tumour-associated antigens, which therefore reduces the likelihood of
XX  escape mutants. Furthermore, immunosuppressive groups that may be present
XX  in whole antigens can be avoided with the use of group-based vaccines.
XX  An additional advantage of an group-based vaccine approach is the ability
XX  to combine selected groups (CTL and HTL), and further, to modify the
XX  composition of the groups, achieving, for example, enhanced
XX  immunogenicity. Accordingly, the immune response can be modulated, as
XX  appropriate, for the target disease. Similar engineering of the response
XX  is not possible with traditional approaches. ABP11501 to ABP25412
XX  represent peptide sequences used in the exemplification of the present
XX  invention.
XX
XX  Sequence 8 AA:
XX
XX  Query Match 79.5%; Score 35; DB 22; Length 8;
XX  Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY  2 KARVLAEA 9
    |||||
DB  1 KARVLAEA 8

```

```

RESULT 11
ABP20832
ID  ABP20832 standard; Peptide: 8 AA.
XX
XX  ABP20832;
XX
XX  15-JUL-2002 (first entry)
XX

```

```

XX  HIV A03 motif gag peptide #495.
XX
XX  HIV; HIV-1; human immunodeficiency virus: env; pol; gag; nef; vpr;
XX  vpu; vif; tat; cytoxic T lymphocyte; CTL; immune response; epitope;
XX  antigen; vaccine; HIV infection; immunisation; virucide.
XX
XX  Human immunodeficiency virus type 1.
XX
XX  WO200124810-A1.
XX
XX  12-APR-2001.
XX
XX  05-OCT-2000; 2000WO-US27766.
XX
XX  05-OCT-1999; 99US-0412863.
XX
XX  (EPI-M-) EPIMUNE INC.
XX
XX  Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX  Baker DM, Cells E, Kubo RT, Grey HM;
XX  WPI: 2001-354887/37.
XX
XX  Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
XX  peptide groups, useful for vaccinating against HIV-1.
XX
XX  Claim 32; Page 296; 448pp; English.
XX
XX  The present invention describes a composition (I) comprising a prepared
XX  human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
XX  sequence selected from 51 defined amino acid sequences (AB125347 to
XX  ABP25397). (I) has virucide activity and can be used in vaccines. (I)
XX  may be used for immunising subjects against HIV-1 infections. The use of
XX  group-based vaccines has several advantages over traditional vaccines,
XX  particularly when compared to the use of whole antigens in vaccine
XX  compositions. There is evidence that the immune response to whole
XX  antigens is directed largely toward variable regions of the antigen,
XX  allowing for immune escape due to mutations. The groups for inclusion in
XX  an group-based vaccine may be selected from conserved regions of viral or
XX  tumour-associated antigens, which therefore reduces the likelihood of
XX  escape mutants. Furthermore, immunosuppressive groups that may be present
XX  in whole antigens can be avoided with the use of group-based vaccines.
XX  An additional advantage of an group-based vaccine approach is the ability
XX  to combine selected groups (CTL and HTL), and further, to modify the
XX  composition of the groups, achieving, for example, enhanced
XX  immunogenicity. Accordingly, the immune response can be modulated, as
XX  appropriate, for the target disease. Similar engineering of the response
XX  is not possible with traditional approaches. ABP11501 to ABP25412
XX  represent peptide sequences used in the exemplification of the present
XX  invention.
XX
XX  Sequence 8 AA:
XX
XX  Query Match 79.5%; Score 35; DB 22; Length 8;
XX  Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY  2 KARVLAEA 9
    |||||
DB  1 KARVLAEA 8

```

```

RESULT 12
ABP12751
ID  ABP12751 standard; Peptide: 9 AA.
XX
XX  ABP12751;
XX
XX  15-JUL-2002 (first entry)
XX
XX  HIV A02 super motif gag peptide #222.
XX

```

KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
 KM vpu; tat; cytoxic T lymphocyte; CTL; immune response; epitope;  
 KM antigen; vaccine; HIV infection; immunisation; virucide.  
 OS Human immunodeficiency virus type 1.  
 XX WO200124810-A1.  
 PN 12-APR-2001.  
 XX 05-OCT-2000; 2000WO-US27766.  
 PF 05-OCT-1999; 99US-0412863.  
 XX (EPIM-) EPIMUNE INC.  
 PA Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Cells E, Kubo RT, Grey HM;  
 XX WPI: 2001-354887/37.  
 DR Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
 XX peptide groups, useful for vaccinating against HIV-1.  
 PT Claim 32: Page 130; 448pp; English.  
 PS The present invention describes a composition (I) comprising a prepared  
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)  
 CC may be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines.  
 CC An additional advantage of an group-based vaccine approach is the ability  
 CC to combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABP1501 to ABP25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention.  
 CC SQ Sequence 9 AA;  
 XX  
 XX Query Match 79.5%; Score 35; DB 22; Length 9;  
 XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KARVLAEA 9  
 XX |||||  
 Db 1 KARVLAEA 8  
 XX  
 XX RESULT 13  
 XX ABL17961  
 XX ID ABL17961 standard; Peptide; 9 AA.  
 XX AC ABL17961;  
 XX 15-JUL-2002 (first entry)  
 XX DE HIV B58 super motif gag peptide #180.  
 XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
 XX vpu; tat; cytoxic T lymphocyte; CTL; immune response; epitope;  
 KW antigen; vaccine; HIV infection; immunisation; virucide.

XX  
 OS Human immunodeficiency virus type 1.  
 XX WO200124810-A1.  
 PN 12-APR-2001.  
 XX 05-OCT-2000; 2000WO-US27766.  
 PF 05-OCT-1999; 99US-0412863.  
 XX (EPIM-) EPIMUNE INC.  
 PA Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Cells E, Kubo RT, Grey HM;  
 XX WPI: 2001-354887/37.  
 DR Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
 XX peptide groups, useful for vaccinating against HIV-1.  
 PT Claim 32: Page 237; 448pp; English.  
 PS The present invention describes a composition (I) comprising a prepared  
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)  
 CC may be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines.  
 CC An additional advantage of an group-based vaccine approach is the ability  
 CC to combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABP1501 to ABP25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention.  
 CC SQ Sequence 9 AA;  
 XX  
 XX Query Match 79.5%; Score 35; DB 22; Length 9;  
 XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KARVLAEA 9  
 XX |||||  
 Db 1 KARVLAEA 8  
 XX  
 XX RESULT 14  
 XX AAE19118  
 XX ID AAE19118 standard; peptide; 10 AA.  
 XX AC AAE19118;  
 XX 21-MAY-2002 (first entry)  
 XX DE Protease peptide.  
 XX DE Luciferase; enzyme; protease; cell growth; apoptosis; therapeutic;  
 KW cell death.  
 XX OS Unidentified.  
 XX MO200206458-A2.

XX 24-JAN-2002.  
PD 17-JUL-2001; 2001WO-US22478.  
XX 19-JUL-2000; 2000US-0619047.  
XX (CHEM-) CHEMICON INT INC.  
XX Long J;  
PI WPI: 2002-195809/25.  
DR  
XX  
XX New purified luciferase polypeptide, useful for identifying protease  
PT activity modulators, comprises recognition site cleavable by protease,  
PT where cleavage of polypeptide results in decreased luciferase activity  
PT  
XX  
XX Claim 63; Page 24; 56pp; English.  
PS  
XX  
XX The invention relates to a purified polypeptide having luciferase  
CC activity and a recognition site specifically cleavable by a protease,  
CC where cleavage results in a decrease in luciferase activity. The  
CC polypeptide comprises a localisation sequence which is linked to the  
CC luciferase polypeptide by the cleavable recognition sequence. The  
CC polypeptide is useful for identifying a protease (Caspase-3) activity  
CC modulator, an inhibitor of apoptosis and for detecting luciferase  
CC activity in a sample. The polypeptide is used for characterising and  
CC identifying cellular processes associated with metabolism, cell growth  
CC and cell death e.g. apoptosis and for measuring protease activity. The  
CC protease measurement methods are useful for characterising, identifying  
CC cellular biochemical pathways as well as identifying diagnostic and  
CC agents for modulating diseases or disorders associated with  
CC biochemical pathways. The polypeptide is also useful as a substrate to  
CC study agents or conditions that cleave the recognition site and for  
CC determining amount of protease in a sample and for monitoring the  
CC activity of protease inside a cell that expresses a recombinant  
CC luciferase. The present sequence is a protease peptide used in the  
CC invention.  
XX  
XX Sequence 10 AA:  
SQ  
Query Match 79.5%; Score 35; DB 23; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 KARVLAEA 9  
DB 1 KARVLAEA 8  
RESULT 15  
AAU75795  
ID AAU75795 standard; Peptide: 10 AA.  
XX  
XX AAU75795;  
AC  
XX  
XX 08-MAY-2002 (first entry)  
DT  
XX  
XX Human immunodeficiency virus Gag protein protease cleavage sequence #2.  
DE  
XX  
XX Gag protein; antiviral; nucleic acid construct; viral protease;  
KW protease recognition site; PRS; viral drug resistance; HIV;  
KW human immunodeficiency virus.  
XX  
XX Human immunodeficiency virus type 1.  
OS  
XX  
XX Key Location/Qualifiers  
FH Cleavage-site 5  
FT /Label- HIV\_protease\_cleavage-site  
XX  
XX W0200210430-A2.  
XX

PD 07-FEB-2002.  
XX 30-JUL-2001; 2001WO-IL00702.  
XX 31-JUL-2000; 2000US-0629969.  
XX (AMID-) AMIDUT LTD.  
XX  
XX Appel E;  
PI WPI: 2002-188753/24.  
DR  
XX  
XX New nucleic acid construct for detecting anti-viral drugs, comprises a  
PT polynucleotide cassette encoding a chimeric polypeptide with two  
PT polypeptide sequences and a protease recognition site cleavable by a  
PT virally encoded protease -  
XX  
XX Example 1; Page 32; 93pp; English.  
PS  
XX  
XX The present invention relates to a new nucleic acid construct that  
CC comprises a polynucleotide cassette encoding a chimeric polypeptide.  
CC The chimeric polypeptide comprises a first polypeptide sequence, a  
CC second polypeptide sequence translationally fused to the first sequence,  
CC and a protease recognition site (PRS) cleavable by a virally encoded  
CC protease, where cleavage of the PRS leads to a detectable signal. The  
CC nucleic acid construct of the invention is useful for uncovering  
CC molecules having antiviral activity or for determining viral drug  
CC resistance and the chimeric peptide is also useful for detecting the  
CC presence of a virus in a cell. Other uses of the invention are detecting  
CC viral encoded protease, found in infected cells or detecting the presence  
CC or absence of viral infection. The nucleic acid construct is useful for  
CC phenotypic testing of human immunodeficiency virus (HIV) drug resistance.  
CC The invention enables screening of molecules in an easy and rapid manner  
CC and allows efficient detection of the presence of viral protease and  
CC viral particles within cells. The construct is specific, sensitive and  
CC lacks background enzymatic activity in the absence of human  
CC immunodeficiency virus protease, and is suitable for detection of  
CC specific viral strain isolates even under low viral load conditions. When  
CC used for phenotypic testing of HIV drug resistance it delivers accurate  
CC results within 24 hours. The present amino acid sequence represents one  
CC of several (AAU75794-AAU75798) HIV Gag protein protease cleavage  
CC sequences that were used in the invention for HIV detection.  
XX  
XX Sequence 10 AA:  
SQ  
Query Match 79.5%; Score 35; DB 23; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 KARVLAEA 9  
DB 1 KARVLAEA 8  
Search completed: March 13, 2003, 18:49:14  
Job time : 27.5455 secs

Examiner's ref.

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:52 ; Search time 8.90909 Seconds  
(without alignments)  
33.026 Million cell updates/sec

Title: US-10-050-200-5  
Perfect score: 44  
Sequence: 1 EKARVLAEAA 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/1aa/PCFUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	79.5	10	2	US-08-792-553-4
2	35	79.5	10	4	US-09-053-941-15
3	35	79.5	10	4	US-09-817-413-15
4	35	79.5	12	4	US-09-147-208-21
5	35	79.5	15	1	US-08-257-5288-76
6	35	79.5	15	1	US-08-460-602A-76
7	35	79.5	15	1	US-08-463-966A-76
8	35	79.5	15	1	US-08-465-217A-76
9	35	79.5	15	2	US-08-464-329A-76
10	35	79.5	15	2	US-08-462-507A-76
11	35	79.5	15	2	US-08-467-881A-76
12	35	79.5	32	1	US-08-257-5288-47
13	35	79.5	32	1	US-08-460-602A-47
14	35	79.5	32	1	US-08-463-966A-47
15	35	79.5	32	1	US-08-465-217A-47
16	35	79.5	32	2	US-08-464-329A-47
17	35	79.5	32	2	US-08-462-507A-47
18	35	79.5	32	2	US-08-467-881A-47
19	35	79.5	34	1	US-08-257-5288-78
20	35	79.5	34	1	US-08-257-5288-77
21	35	79.5	34	1	US-08-460-602A-77
22	35	79.5	34	1	US-08-463-966A-77
23	35	79.5	34	1	US-08-465-217A-77
24	35	79.5	34	1	US-08-462-507A-77
25	35	79.5	34	1	US-08-467-881A-77
26	35	79.5	34	1	US-08-465-217A-78
27	35	79.5	34	2	US-08-464-329A-77

## ALIGNMENTS

28	35	79.5	34	2	US-08-464-329A-78	Sequence 78, Appl
29	35	79.5	34	2	US-08-462-507A-77	Sequence 77, Appl
30	35	79.5	34	2	US-08-462-507A-78	Sequence 78, Appl
31	35	79.5	34	2	US-08-467-881A-77	Sequence 77, Appl
32	35	79.5	34	2	US-08-467-881A-78	Sequence 78, Appl
33	35	79.5	37	4	US-09-621-625A-27	Sequence 27, Appl
34	35	79.5	163	4	US-09-117-217-8	Sequence 8, Appl
35	35	79.5	172	2	US-08-117-881-6	Sequence 6, Appl
36	35	79.5	172	2	US-08-477-081-6	Sequence 6, Appl
37	35	79.5	172	5	PCT-US93-021A2-6	Sequence 6, Appl
38	35	79.5	172	5	US-09-621-625A-25	Sequence 25, Appl
39	35	79.5	275	1	US-08-589-446-4	Sequence 4, Appl
40	35	79.5	275	1	US-08-444-882-4	Sequence 4, Appl
41	35	79.5	275	2	US-08-389-459A-4	Sequence 4, Appl
42	35	79.5	275	3	US-08-987-867A-4	Sequence 4, Appl
43	35	79.5	292	4	US-09-621-625A-2	Sequence 2, Appl
44	35	79.5	292	4	US-09-621-625A-17	Sequence 17, Appl
45	35	79.5	294	4	US-09-370-368-3	Sequence 3, Appl

RESULT 1  
US-08-792-553-4  
Sequence 4, Application US/08792553  
Patent No 4,598,120

## GENERAL INFORMATION:

APPLICANT: Tsien, Roger Y.

TITLE OF INVENTION: Tandem Fluorescent Protein Constructs

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FISH & RICHARDSON P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: California

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/792,553

FILING DATE: 31-JAN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Lisa A. Halle, Ph.D.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07257/041001/UC 96-160-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-678-5070

TELEFAX: 619-678-5099

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-792-553-4

Query Match 79.5% Score 35; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 8; conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEAA  
|||||  
Db 1 KARVLAEAA 8

RESULT 2  
US-09-053-941-15  
; Sequence 15, Application US/09053941  
; Patent No. 6271354  
; GENERAL INFORMATION:  
; APPLICANT: SRINIVASAN, ALGARSAMY  
; APPLICANT: KOPROWSKI, HILARY  
; TITLE OF INVENTION: CHIMERIC VIRAL PROTEINS  
; FILE REFERENCE: Chimeric Viral Proteins  
; CURRENT APPLICATION NUMBER: US/09/053,941  
; EARLIER FILING DATE: 1998-04-02  
; EARLIER APPLICATION NUMBER: 60/043,380  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
; PUBLICATION INFORMATION:  
; JOURNAL: Mech. Enzymol.  
; VOLUME: 38  
; PAGES: 299-  
; DATE: 1974  
US-09-053-941-15

Query Match 79.5%; Score 35; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEA 9  
|||||||  
DB 1 KARVLAEA 8

RESULT 3  
US-09-817-413-15  
; Sequence 15, Application US/09817413  
; Patent No. 6436648  
; GENERAL INFORMATION:  
; APPLICANT: SRINIVASAN, ALGARSAMY  
; APPLICANT: KOPROWSKI, HILARY  
; TITLE OF INVENTION: CHIMERIC VIRAL PROTEINS  
; FILE REFERENCE: Chimeric Viral Proteins  
; CURRENT APPLICATION NUMBER: US/09/817,413  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: 60/043,380  
; PRIOR FILING DATE: 1998-04-03  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
; PUBLICATION INFORMATION:  
; JOURNAL: Mech. Enzymol.  
; VOLUME: 38  
; PAGES: 299-  
; DATE: 1974  
US-09-817-413-15

Query Match 79.5%; Score 35; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEA 9  
|||||||  
DB 1 KARVLAEA 8

RESULT 4  
US-09-147-208-21  
; Sequence 21, Application US/09147208

; Patent No. 6333303  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins  
; NUMBER OF SEQUENCES: 71  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERSKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/147,208  
; FILING DATE: 02-MAR-1999  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rudolph, John R.  
; REGISTRATION NUMBER: 38,003  
; REFERENCE/DOCKET NUMBER: 7841-76  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-147-208-21

Query Match 79.5%; Score 35; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEA 9  
|||||||  
DB 2 KARVLAEA 9

RESULT 5  
US-08-257-528B-76  
; Sequence 76, Application US/08257528B  
; Patent No. 5639854  
; GENERAL INFORMATION:  
; APPLICANT: SIA, Charles D.Y.  
; APPLICANT: CHONG, Pele  
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/257,528B  
; FILING DATE: 09-JUN-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:

NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-336 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1153  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-257-528B-76

Query Match 79.5%; Score 35; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
Db 3 KARVLAEA 10

RESULT 6  
US-08-460-602A-76  
Sequence 76, Application US/08460602A  
Patent No. 5759769  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,602A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-450 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1153  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-602A-76

Query Match 79.5%; Score 35; DB 1; Length 15;

Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
Db 3 KARVLAEA 10

RESULT 7  
US-08-463-966A-76  
Sequence 76, Application US/08463966A  
Patent No. 5795955  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,966A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-487 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1153  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-463-966A-76

Query Match 79.5%; Score 35; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
Db 3 KARVLAEA 10

RESULT 8  
US-08-465-217A-76  
Sequence 76, Application US/08465217A  
Patent No. 5800822  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele

APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Slim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,217A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-486 MIS:jb  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1153  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-465-217A-76

Query Match 79.5%; Score 35; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAAE 9  
DB 3 KARVLAAE 10

RESULT 9  
US-08-464-329A-76  
Sequence 76, Application US/08464329A  
Patent No. 5817754  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Slim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,329A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-449 MIS:jb  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1153  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-464-329A-76

Query Match 79.5%; Score 35; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAAE 9  
DB 3 KARVLAAE 10

RESULT 10  
US-08-462-507A-76  
Sequence 76, Application US/08462507A  
Patent No. 5876731  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Slim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,507A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973



```
REFERENCE/DOCKET NUMBER: 1038-451 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-462-507A-76

Query Match
Best Local Similarity 100.0%; Pred No. 0.55;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9
Db 3 KARVLAEA 10

RESULT 11
US-08-467-881A-76
Sequence 76, Application US/08467881A
Patent No. 5951986
GENERAL INFORMATION:
APPLICANT: SIA, Charles D.Y.
APPLICANT: CHONG, Pele
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,881A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/257,528
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/073,378
FILING DATE: 09-JUN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-488 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-881A-76

Query Match
Best Local Similarity 100.0%; Pred No. 0.55;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2 KARVLAEA 9
Db 3 KARVLAEA 10

RESULT 12
US-08-257-528B-47
Sequence 47, Application US/08257528B
Patent No. 5639854
GENERAL INFORMATION:
APPLICANT: SIA, Charles D.Y.
APPLICANT: CHONG, Pele
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/257,528B
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-336 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-257-528B-47

Query Match
Best Local Similarity 100.0%; Pred No. 1.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9
Db 3 KARVLAEA 10

RESULT 13
US-08-460-602A-47
Sequence 47, Application US/08460602A
Patent No. 5759769
GENERAL INFORMATION:
APPLICANT: SIA, Charles D.Y.
APPLICANT: CHONG, Pele
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,602A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-450 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-602A-47

Query Match 79.5%; Score 35; DB 1; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
Db 3 KARVLAEA 10

RESULT 14  
US-08-463-966A-47  
Sequence 47, Application US/08463966A  
Patent No. 5795955  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michael H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Slim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,966A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378

FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-487 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-463-966A-47

Query Match 79.5%; Score 35; DB 1; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
Db 3 KARVLAEA 10

RESULT 15  
US-08-465-217A-47  
Sequence 47, Application US/08465217A  
Patent No. 5800822  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michael H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Slim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,217A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-486 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-465-217A-47

Query Match 79.5%; Score 35; DB 1; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAA 9  
 |||||  
 DB 3 KARVLAA 10

Search completed: March 13, 2003, 18:53:52  
 Job time : 9.90909 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:45:22 ; Search time 26.9455 seconds

(without alignments)  
99.409 Million cell updates/sec

Title: US-10-050-200-7

Perfect score: 62

Sequence: 1 ERAEQRLKSODL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_21:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp rodent:\*
- 12: sp.virus:\*
- 13: sp.vertebrate:\*
- 14: sp.unclassified:\*
- 15: sp.rvirus:\*
- 16: sp.bacteriaph:\*
- 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	74.2	177	17	Q978R7
2	46	74.2	179	17	Q9HLE1
3	44	71.0	281	13	Q90ZC6
4	44	71.0	282	13	Q90Y46
5	44	71.0	293	13	Q90Y46
6	43	69.4	756	16	Q9RYG4
7	41	66.1	175	4	Q15608
8	41	66.1	177	11	Q920X5
9	41	66.1	190	11	Q9D1S7
10	41	66.1	230	13	Q91809
11	41	66.1	234	6	Q9TS30
12	41	66.1	249	6	Q9TS31
13	41	66.1	250	4	Q15607
14	41	66.1	270	6	Q95MC6
15	41	66.1	271	11	Q9QZ47
16	41	66.1	272	11	Q99189

17	41	66.1	275	6	Q95MC7
18	41	66.1	278	13	Q9DE10
19	41	66.1	275	13	Q13096
20	41	66.1	279	13	Q918R9
21	41	66.1	284	6	Q9TUM9
22	41	66.1	285	4	Q9BUP6
23	41	66.1	287	13	Q57559
24	41	66.1	291	6	Q95MC8
25	41	66.1	291	11	Q64360
26	41	66.1	295	11	Q64377
27	41	66.1	296	6	Q95MC9
28	41	66.1	297	13	Q9DE38
29	41	66.1	301	13	Q9DE11
30	41	66.1	310	11	Q9R563
31	41	66.1	380	13	Q9W7A2
32	39	62.9	218	4	Q9NP46
33	39	62.9	223	13	Q93376
34	39	62.9	228	11	Q92357
35	39	62.9	250	11	Q88795
36	39	62.9	251	4	Q95472
37	39	62.9	261	11	Q9QUP7
38	39	62.9	262	11	Q88346
39	39	62.9	309	11	Q63097
40	39	62.9	322	4	Q961X2
41	39	62.9	331	11	Q8R4T3
42	39	62.9	375	6	Q95KH8
43	39	62.9	418	4	Q9BX75
44	39	62.9	560	4	Q961O5
45	39	62.9	895	5	Q8SSZ4

## ALIGNMENTS

RESULT 1

Q978R7 PRELIMINARY: PRT: 177 AA.

AC Q978R7; 01-OCT-2001 (TREMBLrel. 18, Created)

DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE N-terminal acetyltransferase complex subunit.

GN TY1348 OR TYG1394372.

OS Thermoplasma volcanium.

OC Archaea; Euryarchaeota; Thermoplasmatia; Thermoplasmatales;

OC Thermoplasmataceae; Thermoplasma.

OX NCBI\_TaxID=50339;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GS1 / DSM 4299 / JCM 9571;

RA MEDLINE=20570466; PubMed=11121031;

RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,

RA Kawashima T., Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,

RA Nunoshima T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.,

RT "Archaeal adaptation to higher temperatures revealed by genomic

RT sequence of Thermoplasma volcanium."

RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).

DR EMBL: AP000996; BAB60490.1; -

DR InterPro: IPR00182; GCSNacetyltransf.

DR Pfam: PF00583; Acetyltransf. 1.

KW Transferase; Complete proteome.

SQ SEQUENCE 177 AA; E22EA1621005DB2E CRC64;

Query Match 74.2%; Score 46; DB 17; Length 177;

Best local similarity 75.0%; Pred. No. 1.8;

Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ERAEQRLKSOD 12

DB 49 ERAEQRLKSOD 60

RESULT 2

```

O9HLE1
ID O9HLE1 PRELIMINARY; PRT: 179 AA.
AC O9HLE1;
GN
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein TA0277.
GN TA0277.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Friehman D., Stocker S., Lupas A.N., Baumister W.;
RT "The genome sequence of the thermophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513(2000).
DR EMBL: AL445063; CAC11422.1;
DR InterPro: IPR000182; GCM5acetyltransf.
DR Pfam: PF00583; Acetyltransf. 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 179 AA: 20547 MW: 2812857C836FDC CRC64;

Query Match
Best Local Similarity 74.2%; Score 46; DB 17; Length 179;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ERAEQORLSQD 12
DB 49 ERAEQERIKSLD 60

RESULT 3
O90ZC6 PRELIMINARY; PRT: 281 AA.
ID O90ZC6
AC O90ZC6;
GN
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Low Mr mutant cardiac tropoin T.
OS Meleagris galllopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RA Jin J.-P., Blesladeckl B.;
RT "Mutant turkey cardiac tropoin T."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF274301; AAK58685.1;
DR InterPro: IPR001978; Tropoin.
DR Pfam: PF00992; Tropoin; 1.
SQ SEQUENCE 281 AA: 33651 MW: 79F0B960E7674718 CRC64;

Query Match
Best Local Similarity 71.0%; Score 44; DB 13; Length 281;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLSQ 11
DB 126 ERAEQORIRSE 136

RESULT 4
O90Y46 PRELIMINARY; PRT: 282 AA.
ID O90Y46
AC O90Y46;
GN
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

```

```

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cardiac tropoin T.
GN TNNT2.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Sehner A.J., Weinstein B.M., Walker C., Fishman M.C.,
RA Stainier D.Y.R.;
RT "The zebrafish silent heart mutation disrupts Myofibrillogenesis and
RT Cell-Autonomously Reduces Cardiac Tropoin T Expression."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF282384; AAL06279.1;
DR ZFIN: ZDB-GENE-000626-1; tnnt2.
DR InterPro: IPR001978; Tropoin.
DR Pfam: PF00992; Tropoin; 1.
SQ SEQUENCE 282 AA: 33986 MW: EE7450B0A50B1D25 CRC64;

Query Match
Best Local Similarity 71.0%; Score 44; DB 13; Length 282;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLSQ 11
DB 135 ERAEQORIRSE 145

RESULT 5
O90YN7 PRELIMINARY; PRT: 293 AA.
ID O90YN7
AC O90YN7;
GN
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Cardiac tropoin T isoform.
OS Meleagris galllopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RA Jin J.-P., Blesladeckl B.;
RT "Mutant turkey cardiac tropoin T."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY005139; AAG23714.1;
DR InterPro: IPR001978; Tropoin.
DR Pfam: PF00992; Tropoin; 1.
SQ SEQUENCE 293 AA: 34892 MW: 7D7B5EFCAE16C2B2 CRC64;

Query Match
Best Local Similarity 71.0%; Score 44; DB 13; Length 293;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLSQ 11
DB 138 ERAEQORIRSE 148

RESULT 6
O9RYG4 PRELIMINARY; PRT: 756 AA.
ID O9RYG4
AC O9RYG4;
GN
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Methyl-accepting chemotaxis-related protein.
GN DRA0353.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.

```

OX NCBI\_TaxID-1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RL;  
 RX MEDLINE-20036886; PubMed-10567266;  
 RA White O., Eissen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffit K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamatheyan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 RT radiodurans RL";  
 RL Science 286:1571-1577(1999).  
 DR EMBL: AE001863; AAF12432.1; -;  
 DR HSSP: P02942; 1Q07.  
 DR TIGR: DRA0353; -;  
 DR InterPro: IPR004089; Chmtaxis\_transd.  
 DR InterPro: IPR003660; HAMP.  
 DR Pfam: PF00672; HAMP; 1.  
 DR Pfam: PF00015; MCPsignal; 1.  
 DR SMART: SM00304; HAMP; 2.  
 DR SMART: SM00283; MA; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 756 AA; 81011 MW; 8853960403E06204 CRC64;

Query Match 69.4%; Score 43; DB 16; Length 756;  
 Best Local Similarity 69.2%; Pred. No. 25;  
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ERAEQORLKSQDL 13  
 DB 416 ERVEQERLESQRL 428

RESULT 7  
 ID 015608 PRELIMINARY; PRT; 175 AA.  
 AC 015608;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Troponin T (Fragment).  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID-9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-96129582; PubMed-8576938;  
 RA Townsend P.J., Barton P.J.R., Yacoub M.H., Farza H.;  
 RT "Molecular cloning of human cardiac troponin T isoforms: expression in  
 RT developing and failing heart";  
 RL J. Mol. Cell. Cardiol. 27:2223-2236(1995).  
 DR EMBL: X79859; CA56239.1; -;  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 FT NON\_TER 175  
 SQ SEQUENCE 175 AA; 20918 MW; B270D28FB067ED02 CRC64;  
 Query Match 66.1%; Score 41; DB 4; Length 175;  
 Best Local Similarity 63.6%; Pred. No. 14;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

AC 0920X5;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Troponin T (Fragment).  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.  
 OX NCBI\_TaxID-10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99345814; PubMed-10409222;  
 RA Preisig-Muller R., Mederos Y Schnitzler M., Derst C., Daut J.;  
 RT "Separation of cardiomyocytes and coronary endothelial cells for cell-  
 RT specific RT-PCR";  
 RL Am. J. Physiol. 277:H413-H416(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Preisig-Muller R., Derst C., Mederos Y Schnitzler M., Daut J.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF099071; AAD16291.1; -;  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 FT NON\_TER 177  
 FT NON\_TER 177  
 SQ SEQUENCE 177 AA; 21787 MW; 863673ABBD634223 CRC64;  
 Query Match 66.1%; Score 41; DB 11; Length 177;  
 Best Local Similarity 63.6%; Pred. No. 14;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLKSQ 11  
 DB 37 ERAEQORIRNE 47

RESULT 9  
 ID 09D1S7 PRELIMINARY; PRT; 190 AA.  
 AC 09D1S7;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE Troponin T3, skeletal, fast.  
 GN TNNT3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID-10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6J; TISSUE-EMBRYO;  
 RX MEDLINE-21085660; PubMed-11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirni L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 Nature 409:685-690(2001).

DR EMBL: AK003147; BAB22603.1; -.  
 DR MGD: MGI:109550; Tnnt3.  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 DR SEQUENCE 190 AA; 22497 MW; E4B84679D4DBE0C9 CRC64;

Query Match 66.1%; Score 41; DB 11; Length 190;  
 Best Local Similarity 63.6%; Pred. No. 15;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLKSQ 11  
 DB 34 ERAEQORIRAE 44

## RESULT 10

ID 091809 PRELIMINARY: PRT: 230 AA.  
 AC 091809;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Fast skeletal muscle troponin T.  
 GN TNNI.  
 OS Brachydanio rerio (Zebrafish) (Zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_Taxid=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20458823; PubMed=11002340;  
 RA Xu Y., He J., Wang X., Lim T.M., Gong Z.;  
 RT "Asynchronous activation of 10 muscle-specific protein (NSP) genes  
 RL during zebrafish somitogenesis.";  
 RL Dev. Dyn. 219:201-215(2000).  
 DR EMBL: AF180889; AAF78472.1; -.  
 DR ZFIN: ZDB-GENE-000322-3; tnni.  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 DR SEQUENCE 230 AA; 27786 MW; 37729A66290252DD CRC64;

Query Match 66.1%; Score 41; DB 13; Length 230;  
 Best Local Similarity 63.6%; Pred. No. 18;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLKSQ 11  
 DB 79 ERAEQORIRTE 89

## RESULT 11

ID 09TS30 PRELIMINARY: PRT: 234 AA.  
 AC 09TS30;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE Troponin T alpha isoform.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_Taxid=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93054628; PubMed=1429653;  
 RA Pan B.S., Potter J.D.;  
 RT "Two genetically expressed troponin T fragments representing alpha and  
 RL beta isoforms exhibit functional differences.";  
 RL J. Biol. Chem. 267:23052-23056(1992).  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 DR SEQUENCE 234 AA; 27809 MW; 1AFA69100C5A97A CRC64;

Query Match 66.1%; Score 41; DB 6; Length 234;  
 Best Local Similarity 63.6%; Pred. No. 18;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLKSQ 11  
 DB 78 ERAEQORIRAE 88

## RESULT 12

ID 09TS31 PRELIMINARY: PRT: 249 AA.  
 AC 09TS31;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE Troponin T beta isoform.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_Taxid=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93054628; PubMed=1429653;  
 RA Pan B.S., Potter J.D.;  
 RT "Two genetically expressed troponin T fragments representing alpha and  
 RL beta isoforms exhibit functional differences.";  
 RL J. Biol. Chem. 267:23052-23056(1992).  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 DR SEQUENCE 249 AA; 29683 MW; ABEA7AC4E86FCCF CRC64;

Query Match 66.1%; Score 41; DB 6; Length 249;  
 Best Local Similarity 63.6%; Pred. No. 19;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLKSQ 11  
 DB 93 ERAEQORIRAE 103

## RESULT 13

ID 015607 PRELIMINARY: PRT: 250 AA.  
 AC 015607;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Cardiac troponin T (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEART MUSCLE;  
 RX MEDLINE=96129582; PubMed=8576938;  
 RA Townsend P.J., Barton P.J.R., Yacoub M.H., Farza H.;  
 RT "Molecular cloning of human cardiac troponin T isoforms: expression in  
 RL developing and failing heart.";  
 RL J. Mol. Cell. Cardiol. 27:2223-2236(1995).  
 DR EMBL: X79861; CAA56240.1; -.  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 DR NON\_TER 250  
 DR SEQUENCE 250 AA; 30151 MW; 975F687EB21B33D6 CRC64;

Query Match 66.1%; Score 41; DB 4; Length 250;  
 Best Local Similarity 63.6%; Pred. No. 20;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLKSQ 11



```

Db      128 ERAEQORIRNE 138

RESULT 14
095MC6 PRELIMINARY; PRT; 270 AA.
AC 095MC6
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
OS Tropoin T splicing mutant isoform 2.
OC Cans familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN (1)
RP SEQUENCE FROM N.A.
RA Jin J.-P.;
RT "Dog cardiac tropoin T splicing variants relating to
RT cardiomyopathy."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005143; AAG23718.1; -
DR InterPro: IPR001978; Tropoin.
DR Pfam: PF00992; Tropoin; 1.
SQ SEQUENCE 270 AA; 32339 MW; 316105372CACG80D CRC64;

Query Match 66.1%; Score 41; DB 6; Length 270;
Best Local Similarity 63.6%; Pred. No. 21;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERAEQORIRNSQ 11
Db 116 ERAEQORIRNE 126

RESULT 15
090247 PRELIMINARY; PRT; 271 AA.
AC 090247; 090246; 035575; 035576; 035577; 035578; 035579; 035580;
AC 035581; 035582; 035583; 035584; 035585; P97456;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-JUN-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Tropoin T, fast skeletal muscle isoforms.
GN TNNT3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=129/SVJ; TISSUE=SKELETAL MUSCLE;
RX MEDLINE=97390134; PubMed=9249073;
RA Wang J., Jin J.-P.;
RT "Primary structure and developmental acidic to basic transition of 13
RT alternatively spliced mouse fast skeletal muscle tropoin T
RT isoforms."
RL Gene 193:105-114(1997).
RN (12)
RP SEQUENCE FROM N.A. (ISOFORM B3E17).
RX MEDLINE=97262102; PubMed=9107680;
RA Koch A., Juan T.S.-C., Jenkins N.A., Gilbert D.J., Copeland N.G.,
RA McNelece I.R., Fletcher F.A.;
RT "cDNA cloning and chromosomal mapping of mouse fast skeletal muscle
RT tropoin T."
RL Mamm. Genome 8:346-348(1997).
CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF
CC TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONFERS
CC CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY
CC (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 13 ISOFORMS: A1E17 (SHOWN HERE),
CC A2E17, A3E17, A4E17, A5E17, A6E17, B1E16, B2E17, B2E16,

```

```

CC B3E17/PTROPT, B3E16, B4E17 AND B4E16; ARE PRODUCED BY ALTERNATIVE
CC SPLICING. ISOFORMS A1E17, A2E17, A3E17, A4E17, A5E17 AND A6E17 ARE
CC ACIDIC WHILE ISOFORMS B1E16, B2E17, B2E16, B3E17, B3E16, B4E17 AND
CC B4E16 ARE BASIC.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN SKELETAL MUSCLE.
CC -1- DEVELOPMENTAL STAGE: EMBRYO AND NEONATE EXPRESS PREDOMINANTLY
CC ACIDIC ISOFORMS WHILE ALL ADULT ISOFORMS ARE BASIC. ONLY ONE
CC ISOFORM, B2E17, IS FOUND IN BOTH EMBRYO/NEONATE AND ADULT. A
CC TRANSITION FROM HIGH TO LOW MOLECULAR WEIGHT ISOFORMS IS ALSO SEEN
CC DURING POSTNATAL DEVELOPMENT.
CC EMBL; L49466; AAB67290.1; -
CC DR EMBL; L49467; AAF01502.1; -
CC DR EMBL; L49468; AAF01503.1; -
CC DR EMBL; L49470; AAB67291.1; -
CC DR EMBL; L49471; AAB67292.1; -
CC DR EMBL; L49472; AAB67293.1; -
CC DR EMBL; L48988; AAB67283.1; -
CC DR EMBL; L48989; AAB67284.1; -
CC DR EMBL; L48990; AAB67285.1; -
CC DR EMBL; L48991; AAB67286.1; -
CC DR EMBL; L48992; AAB67287.1; -
CC DR EMBL; L48993; AAB67288.1; -
CC DR EMBL; L49018; AAB67289.1; -
CC DR EMBL; U77779; AAB39743.1; -
CC DR MOD; MG1:109550; Tnnt3.
CC DR InterPro: IPR001978; Tropoin.
CC DR Pfam: PF00992; Tropoin; 1.
CC KW Muscle protein; Alternative splicing; Multigene family;
CC KW Phosphorylation.
CC FT INIT_MET 0 0
CC FT MOD_RES 1 1
CC FT VARSPPLIC 10 15
CC FT VARSPPLIC 22 30
CC FT VARSPPLIC 22 48
CC FT VARSPPLIC 27 30
CC FT VARSPPLIC 31 48
CC FT VARSPPLIC 36 48
CC FT VARSPPLIC 242 254
CC FT CONFLICT 130 130
CC FT CONFLICT 251 251
CC FT SEQUENCE 271 AA; 32109 MW; D2CB544B8C7953F CRC64;
CC SQ SEQUENCE 271 AA; 32109 MW; 316105372CACG80D CRC64;

QY 1 ERAEQORIRNSQ 11
Db 115 ERAEQORIRAE 125

Query Match 66.1%; Score 41; DB 11; Length 271;
Best Local Similarity 63.6%; Pred. No. 21;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Search completed: March 13, 2003, 18:51:59
Job time : 28.9455 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:40:28 ; Search time 6.85455 Seconds

(without alignments)  
78.662 Million cell updates/sec

Title: US-10-050-200-7

Perfect score: 62

Sequence: 1 ERAEQORLKSQDL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40.\*

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	91.9	321	1 FCE2_HUMAN	P06734 homo sapien
2	44	71.0	301	1 TRT2_CHICK	P02642 gallus gall
3	41	66.1	252	1 TRT3_COTJA	P06398 coturnix co
4	41	66.1	257	1 TRT3_HUMAN	P45378 homo sapien
5	41	66.1	258	1 TRT3_RAT	P09799 rattus norv
6	41	66.1	262	1 TRT3_CHICK	P12620 gallus gall
7	41	66.1	278	1 TRT3_RABIT	P02641 oryctolagus
8	41	66.1	284	1 TRT2_BOVIN	P13789 bos taurus
9	41	66.1	286	1 TRT2_SHEEP	P50751 ovis aries
10	41	66.1	297	1 TRT2_HUMAN	P45379 homo sapien
11	41	66.1	298	1 TRT2_RAT	P50753 rattus norv
12	41	66.1	300	1 TRT2_MOUSE	P50752 mus musculu
13	40	64.5	248	1 TPM4_RAT	P09495 rattus norv
14	39	62.9	190	1 GRPE_STRPY	P09498 streptococ
15	39	62.9	184	1 TRT1_HUMAN	P13805 homo sapien
16	39	62.9	184	1 XMS2_DROME	P09399 drosophila
17	39	62.9	1691	1 POLN_HEVME	P03495 hepatitis e
18	38	61.3	157	1 HMAA_SCHGR	P29536 schistocerc
19	38	61.3	911	1 CAPA_MOUSE	P09470 mus musculu
20	38	61.3	938	1 CAPA_HUMAN	P13111 homo sapien
21	38	61.3	1270	1 ITN1_XENLA	P42287 xenopus lae
22	37	59.7	312	1 YMBX_YEAST	P00499 saccharomyc
23	37	59.7	439	1 CUB8_MOUSE	P58467 mus musculu
24	37	59.7	2175	1 HMCU_DROME	P10180 drosophila
25	36.5	58.9	513	1 YMMB_YEAST	P003104 saccharomyc
26	36	58.1	97	1 HMAA_MANSE	P026430 manduca sex
27	36	58.1	331	1 FCE2_HUMAN	P02693 mus musculu
28	36	58.1	386	1 NESG_HUMAN	P09166 homo sapien
29	36	58.1	1171	1 DIA3_MOUSE	P092207 mus musculu
30	36	58.1	1971	1 MC3A_MOUSE	P09039 mus musculu
31	36	58.1	1980	1 MC3A_HUMAN	P06033 homo sapien
32	35	56.5	203	1 L2CC_DROME	P24156 drosophila
33	35	56.5	588	1 PR28_YEAST	P23394 saccharomyc

34	35	56.5	590	1 HMAA_DROME	P29555 drosophila
35	35	56.5	649	1 P137_HUMAN	P14444 homo sapien
36	35	56.5	652	1 DREB_CHICK	P18302 gallus gall
37	35	56.5	656	1 P137_MOUSE	P06885 mus musculu
38	35	56.5	732	1 CADL_CHICK	P33145 mus musculu
39	35	56.5	1597	1 RLRL_YEAST	P53552 saccharomyc
40	35	56.5	1790	1 SEPA_EMENT	P78621 emerficella
41	35	54.8	349	1 DIA2_MOUSE	P70566 mus musculu
42	34	54.8	388	1 PAM_STRPY	P49054 streptococ
43	34	54.8	413	1 FL1_TOBAC	P40504 nicotiana t
44	34	54.8	416	1 FL2_TOBAC	P40505 nicotiana t
45	34	54.8	506	1 Y213_MYCPN	P75478 mycoplasma

## ALIGNMENTS

RESULT 1	FCE2_HUMAN	STANDARD;	PRT;	321 AA.
ID	P06734;			
AC	01-JAN-1988 (Rel. 06, Created)			
DT	01-JAN-1988 (Rel. 06, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Low affinity immunoglobulin epsilon FC receptor (Lymphocyte Ige			
DE	receptor) (FC-epsilon-RII) (CD23) (BLAST-2) (Immunoglobulin E-binding			
DE	factor).			
GN	FCE2 OR IGEBF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=87118255; PubMed=2949326;			
RA	Ikuta K., Takami M., Kim C.W., Honjo T., Miyoshi T., Tagaya Y.,			
RA	Kawabe T., Yodoi J.;			
RT	"Human lymphocyte FC receptor for Ige: sequence homology of its			
RT	cloned cDNA with animal lectins";			
RL	Proc. Natl. Acad. Sci. U.S.A. 84:819-823(1987).			
RN	[2]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=87051737; PubMed=2877743;			
RA	Kikutan H., Inui S., Sato R., Barsumian E.L., Owaki H.,			
RA	Yamasaki K., Kaisho T., Uchiyashiki N., Hardy R.R., Hirano T.,			
RA	Tsunasawa S., Sakiyama F., Suenura M., Kishimoto T.;			
RT	"Molecular structure of human lymphocyte receptor for Immunoglobulin			
RT	E.";			
RL	Cell 47:657-665(1986).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87218454; PubMed=3034567;			
RA	Liedin C., Hofstetter H., Sarfati M., Levy C.A., Suter U., Alaimo D.,			
RA	Klichner E., Frost H., Delespess G.;			
RT	"Cloning and expression of the cDNA coding for a human Lymphocyte Ige			
RT	receptor.";			
RL	EMBO J. 6:109-114(1987).			
RN	[4]			
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.			
RX	MEDLINE=93038513; PubMed=1417742;			
RA	Rose K., Turcatti G., Graber P., Pochon S., Reganey P.-O.,			
RA	Unsen K.U., Magnenat E., Audomney N., Bonnefoy J.-T.;			
RT	"Partial characterization of natural and recombinant human soluble			
RT	CD23.";			
RL	Biochem. J. 286:819-824(1992).			
RN	[5]			
RP	ALTERNATIVE SPLICING.			
RX	MEDLINE=9028672; PubMed=2972386;			
RA	Yokota A., Kikutan H., Tanaka T., Sato R., Barsumian E.L.,			
RA	Suenura M., Kishimoto T.;			
RT	"Two species of human FC epsilon receptor II (Fc epsilon RII/CD23):			
RT	tissue-specific and IL-4-specific regulation of gene expression.";			
RL	Cell 55:611-618(1988).			

RP	[6]	
RQ	3D-STRUCTURE MODELING OF LECTIN DOMAIN.	
RA	MEDLINE=94191542; PubMed=8142907;	
RB	Padian E.A., Helm B.A.;	
RC	"Modeling of the lectin-homology domains of the human and murine low-	
RD	affinity Fc epsilon receptor (Fc epsilon RI/CD23).";	
RE	Receptor 3:325-341(1993).	
RF	(7)	
RG	3D-STRUCTURE MODELING OF 173-285.	
RH	MEDLINE=96276216; PubMed=8745401;	
RI	Bajorath J., Aruffo A.;	
RJ	"Structure-based modeling of the ligand binding domain of the human	
RK	cell surface receptor CD23 and comparison of two independently	
RL	derived molecular models.";	
RM	Protein Sci. 5:240-247(1996).	
RN	-1- FUNCTION: THIS RECEPTOR HAS ESSENTIAL ROLES IN THE REGULATION	
RO	OF ICE PRODUCTION AND IN THE DIFFERENTIATION OF B-CELLS (IT IS	
RP	A B-CELL-SPECIFIC ANTIGEN).	
RQ	-1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS	
RS	A SOLUBLE EXCRETED FORM.	
RT	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE	
RU	PRODUCED BY ALTERNATIVE SPLICING.	
RV	-1- PTM: N- AND O-GLYCOSYLATED.	
RW	-1- MISCELLANEOUS: THERE ARE TWO KINDS OF FC RECEPTORS FOR ICE, WHICH	
RX	DIFFER IN BOTH STRUCTURE AND FUNCTION: HIGH AFFINITY RECEPTORS ON	
RY	BASOPHILS AND MAST CELLS AND LOW AFFINITY RECEPTORS ON LYMPHOCYTES	
RZ	AND MONOCYTES.	
SA	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.	
SB	-1- DATABASE: NAME=PROV; NOTE=CD guide CD23 entry;	
SC	WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd23.htm".	
SD	-----	
SE	This SWISS-PROT entry is copyright. It is produced through a collaboration	
SF	CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
SG	the European Bioinformatics Institute. There are no restrictions on its	
SH	use by non-profit institutions as long as its content is in no way	
SI	modified and this statement is not removed. Usage by and for commercial	
SJ	entities requires a license agreement (see http://www.isdb.ch/announce/	
SK	or send an email to lisense@isdb.slb.ch).	
SL	-----	
SM	EMBL; M15059; AAA52434.1; -;	
SN	EMBL; M14766; AAA52435.1; -;	
SO	EMBL; X04772; CAA28465.1; -;	
SP	EMBL; M23562; AAA52433.1; -;	
SQ	PIR; A26067; LMHUER.	
SR	PIR; A26164; A26164.	
SS	PIR; A26589; A26589.	
ST	PIR; A31924; A31924.	
SV	PIR; J10132; J10132.	
SW	PDB; 1HLI; 31-JAN-94.	
SX	PDB; 1KJE; 03-APR-96.	
SY	Genew; HGNC:3612; FCER2.	
SZ	MIM; 151445; -;	
T0	InterPro: IPR001304; Lectin_C.	
T1	pfam; PF00059; lectin_c_1.	
T2	SMART; SM0034; CLECT; 1.	
T3	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.	
T4	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.	
T5	KW Ig-binding protein; Transmembrane; Glycoprotein; Receptor; B-cell;	
T6	Repeat; Lectin; Signal-anchor; Alternative splicing; 3D-structure.	
T7	CHAIN 1 321	
T8	CHAIN 150 321	
T9	DOMAIN 1 21	
TA	TRANSMEM 22 47	
TB		
TC	DOMAIN 48 321	
TD	DOMAIN 162 284	
TE	SITE 149 150	
TF	REPEAT 69 89	
TG	REPEAT 90 110	
TH	REPEAT 111 131	
TI	DISULFID 160 288	
TJ	DISULFID 163 174	
TK	DISULFID 191 282	
TL		
TM		
TN		
TO		
TP		
TQ		
TR		
TS		
TT		
TU		
TV		
TW		
TX		
TY		
TZ		
UA		
UB		
UC		
UD		
UE		
UF		
UG		
UH		
UI		
UJ		
UK		
UL		
UM		
UN		
UO		
UP		
UQ		
UR		
US		
UT		
UU		
UV		
UW		
UX		
UY		
UZ		
V0		
V1		
V2		
V3		
V4		
V5		
V6		
V7		
V8		
V9		
WA		
WB		
WC		
WD		
WE		
WF		
WG		
WH		
WI		
WJ		
WK		
WL		
WM		
WN		
WO		
WP		
WQ		
WR		
WS		
WT		
WU		
WW		
WX		
WY		
WZ		
X0		
X1		
X2		
X3		
X4		
X5		
X6		
X7		
X8		
X9		
YA	</	

FT	DISEASED	259	273	
FT	CARBOHYD	63	63	N-LINKED (GLCNAC.) (POTENTIAL).
FT	VARSPLIC	1	7	MEROYS -> MNPPSQ (IN ISOFORM B).
SO	SEQUENCE	269	269	N->T (IN REF. 3).
		321 AA;	36468 MM;	F6B708CDE6515B87 CRC64;
	Query Match		91.9%;	Score 57; DB 1; Length 321;
	Best Local Similarity	100.0%;	Pred. No.	0.0084;
	Matches 12;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY	2 RAEOORLKSDQL L3 			
Db	96 RAEOORLKSDQL 107			
	RESULT 2			
ID	TPT2_CHICK	STANDARD;	PRT:	301 AA.
AC	P02642;			
DT	21-JUL-1986 (Rel. 01, Created)			
DR	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Troponin T, cardiac muscle isoforms (TrnC).			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Archosauromia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OX	Gallus.			
RN	NCBI_TaxID=9031;			
RA	[1]			
RX	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RY	MEDLINE=65289327; PubMed=2893302;			
SA	Cooper T.A., Orchard C.P.;			
RT	"A single cardiac troponin T gene generates embryonic and adult			
RL	isoforms via developmentally regulated alternative splicing.";			
JL	J. Biol. Chem. 260:11140-11148(1985).			
RN	[2]			
RP	SEQUENCE OF 67-301 FROM N.A.			
RX	MEDLINE=85065747; PubMed=6095446;			
RA	Cooper T.A., Orchard C.P.;			
RT	"A single troponin T gene regulated by different programs in cardiac			
RL	and skeletal muscle development.";			
SC	Science 226:979-982(1984)			
CC	-I- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF			
CC	TROPOMIN. THE THIN FILAMENT REGULATORY COMPLEX WHICH CONFERS			
CC	CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.			
CC	-I- ALTERNATIVE PRODUCTS: 2 Isoforms; 1/Embryonic form (shown here)			
CC	and 2/Adult form: are produced by alternative splicing.			
CC	-I- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.			
CC	-----			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isdb.ch/announce/			
CC	or send an email to licens@isdb.slb.ch).			
CC	-----			
CC	EMBL: M10013; AAA4909.1; "			
DR	EMBL: K02263; AAA49098.1; "-			
PIR	A03086; TRCHTC.			
PIR	A25373; A25373.			
DR	IInterPro: IPR001978; Troponin.			
DR	Pfam: PF00992; Troponin; 1.			
KW	Muscle protein; Alternative splicing; Multigene family;			
RK	Phosphorylation.			
FT	INT_MET	0		BY SIMILARTY
FT	MOD_RES	1		PHOSPHORYLTATION (BY CK2)
FT	VARSPLIC	22	31	(BY SIMILARTY).
FT	SEQUENCE	301 AA;	35854 MM;	MISSING (IN ISOFORM 2).
SQ				F85CEIA4F07DD94 CRC64;
	Query Match		71.0%;	Score 44; DB 1; Length 301;
	Best Local Similarity	72.7%;	Pred. No.	1.4;

Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ERAEOORLKSQ 11  
 |||||  
 Db 146 ERAEOORIRSE 156

## RESULT 3

TRT3\_COTJA  
 ID TRT3\_COTJA STANDARD; PRT; 252 AA.  
 AC P06398; P06397;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Troponin T, fast skeletal muscle isoforms.  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix.  
 OX NCBI\_TaxID=93934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89308680; PubMed=2745456;  
 RA Bucher E.A., de la Brousse F.C., Emerson C.P. Jr.;  
 RT "Developmental and muscle-specific regulation of avian fast skeletal  
 troponin T isoform expression by mRNA splicing.";  
 RL J. Biol. Chem. 264:12482-12491(1989).  
 RN [2]  
 RP SEQUENCE OF 108-252 FROM N.A.  
 RX MEDLINE=86033836; PubMed=3840482;  
 RA Hastings K.E.M., Bucher E.A., Emerson C.P. Jr.;  
 RT "Generation of troponin T isoforms by alternative RNA splicing in  
 avian skeletal muscle. Conserved and divergent features in birds and  
 mammals.";  
 RL J. Biol. Chem. 260:13699-13703(1985).

CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF  
 TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONTERS  
 CC CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST THREE ISOFORMS ARE PRODUCED BY  
 CC ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M26600; AAA49506.1; -  
 CC EMBL; M26599; AAA49505.1; -  
 CC EMBL; M11684; AAA49504.1; -  
 CC EMBL; M11684; AAA49503.1; -  
 CC PIR; A03085; TPQJ72.  
 CC PIR; A03084; TPQJ71.  
 CC InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 KW Muscle protein; Alternative splicing; Multigene family;  
 KW Phosphorylation.  
 FT INT\_MET 0 0 BY SIMILARITY.  
 FT MOD\_RES 1 1 PHOSPHORYLATION (BY CK2)  
 FT VARSPLIC 27 30 (BY SIMILARITY).  
 FT VARSPLIC 225 237 MISSING (IN AN ISOFORM).  
 FT LITRCLDLSKE -> VILNRIDQAKH (IN AN  
 FT ISOFORM).  
 SQ SEQUENCE 252 AA; 29967 MW; 67FED1A595C3C997 CRC64;

Query Match 66.1%; Score 41; DB 1; Length 252;  
 Best Local Similarity 63.6%; Pred. No. 3.9;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 98 ERAEOORIRAE 108  
 |||||

## RESULT 4

TRT3\_HUMAN  
 ID TRT3\_HUMAN STANDARD; PRT; 257 AA.  
 AC P45378;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Troponin T, fast skeletal muscle isoform beta (Beta TNNT3).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal skeletal muscle;  
 RX MEDLINE=94226695; PubMed=8172653;  
 RA Wu Q.-L., Jha P.K., Raychowdhury M.K., Du Y., Leavis P.C.,  
 RA Sarkar S.;  
 RT "Isolation and characterization of human fast skeletal beta troponin  
 T cDNA: comparative sequence analysis of isoforms and insight into  
 the evolution of members of a multigene family.";  
 RL DNA Cell Biol. 13:217-233(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Stefancsik R., Mao C., Randall J., Jha P.K., Sarkar S.;  
 RT "Genomic structure of the fast skeletal troponin T gene (TNNT3).";  
 RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF  
 TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONTERS  
 CC CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.  
 CC -1- TISSUE SPECIFICITY: IN FETAL AND ADULT FAST SKELETAL MUSCLES, WITH  
 CC A HIGHER LEVEL EXPRESSION IN FETAL THAN IN ADULT MUSCLE.  
 CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M21984; AAA36777.1; -  
 CC EMBL; AF026276; AAF21629.1; -  
 CC Genew; HGNC:11950; TNNT3.  
 DR MIM: 600692;  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 KW Muscle protein; Multigene family; Phosphorylation.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT MOD\_RES 1 1 PHOSPHORYLATION (BY CK2)  
 FT VARSPLIC 257 AA; 30465 MW; F433B3626BE597D9 CRC64;

Query Match 66.1%; Score 41; DB 1; Length 257;  
 Best Local Similarity 63.6%; Pred. No. 4;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Query Match 66.1%; Score 41; DB 1; Length 252;  
 Best Local Similarity 63.6%; Pred. No. 3.9;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

DT 01-FEB-1991 (Rel. 17, last sequence update)  
 DT 15-JUN-2002 (Rel. 41, last annotation update)  
 DE Troponin T, fast skeletal muscle isoforms beta/alpha (Beta/alpha  
 Tntf).  
 CN TINT3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;  
 ON NCBI\_TaxId=10116;  
 RX MEDLINE=66281691; PubMed=3735424;  
 RA Breitbart R.E., Nadal-Ginard B.;  
 RT "Complete nucleotide sequence of the fast skeletal troponin T gene.  
 RT Alternatively spliced exons exhibit unusual interspecies  
 RT divergence.";  
 RL J. Mol. Biol. 188:313-324(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85201690; PubMed=2986851;  
 RA Breitbart R.E., Nguyen H.T., Medford R.M., Destree A.T., Mahdavi V.,  
 RA Nadal-Ginard B.;  
 RT "Intricate combinatorial patterns of exon splicing generate multiple  
 RT regulated troponin T isoforms from a single gene.";  
 RL Cell 41:67-82(1985).  
 CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF  
 CC CALCULUM-SENSITIVITY TO STRIATED MUSCLE ACTINOMYOSIN ATPASE ACTIVITY.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY THE  
 CC USE OF MUTUALLY EXCLUSIVE EXONS.  
 CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M15202; AAA96441.1; -  
 DR EMBL; M15202; AAA96446.1; -  
 DR EMBL; M15202; AAA96464.1; -  
 DR EMBL; M15202; AAA96454.1; -  
 DR EMBL; M15202; AAA96456.1; -  
 DR EMBL; M15202; AAA96463.1; -  
 DR EMBL; M15202; AAA96467.1; -  
 DR EMBL; M15202; AAA96471.1; -  
 DR EMBL; M15202; AAA96475.1; -  
 DR EMBL; M15202; AAA96476.1; -  
 DR EMBL; M15202; AAA96447.1; -  
 DR EMBL; M15202; AAA96442.1; -  
 DR EMBL; M15202; AAA96461.1; -  
 DR EMBL; M15202; AAA96443.1; -  
 DR EMBL; M15202; AAA96465.1; -  
 DR EMBL; M15202; AAA96481.1; -  
 DR EMBL; M15202; AAA96452.1; -  
 DR EMBL; M15202; AAA96468.1; -  
 DR EMBL; M15202; AAA96445.1; -  
 DR EMBL; M15202; AAA96455.1; -  
 DR EMBL; M15202; AAA96457.1; -  
 DR EMBL; M15202; AAA96458.1; -  
 DR EMBL; M15202; AAA96459.1; -  
 DR EMBL; M15202; AAA96440.1; -  
 DR EMBL; M15202; AAA96448.1; -  
 DR EMBL; M15202; AAA96483.1; -  
 DR EMBL; M15202; AAA96449.1; -  
 DR EMBL; M15202; AAA96460.1; -  
 DR EMBL; M15202; AAA96450.1; -  
 DR EMBL; M15202; AAA96462.1; -  
 DR EMBL; M15202; AAA96451.1; -  
 DR EMBL; M15202; AAA96482.1; -

DR EMBL; M15202; AAA96469.1; -  
 DR EMBL; M15202; AAA96453.1; -  
 DR EMBL; M15202; AAA96466.1; -  
 DR EMBL; M15202; AAA96470.1; -  
 DR EMBL; M15202; AAA96444.1; -  
 DR EMBL; M15202; AAA96472.1; -  
 DR EMBL; M15202; AAA96473.1; -  
 DR EMBL; M15202; AAA96474.1; -  
 DR EMBL; M15202; AAA96477.1; -  
 DR EMBL; M15202; AAA96478.1; -  
 DR EMBL; M15202; AAA96479.1; -  
 DR EMBL; M15202; AAA96480.1; -  
 DR PIR; A24824; A24824.  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 KW Muscle protein; Alternative splicing; Multigene family;  
 KW Phosphorylation.  
 FT INIT\_MET 0  
 FT MOD\_RES 1  
 FT VARSPPLIC 9 14 PHOSPHORYLATION (BY CK2)  
 FT VARSPPLIC 15 20 (BY SIMILARITY).  
 FT VARSPPLIC 21 25 MISSING (IN SOME CLASS I ISOFORMS).  
 FT VARSPPLIC 27 30 MISSING (IN SOME CLASS I ISOFORMS).  
 FT VARSPPLIC 31 34 MISSING (IN CLASS II ISOFORMS).  
 FT VARSPPLIC 31 34 MISSING (IN CLASS III ISOFORMS).  
 FT VARSPPLIC 229 241 MISSING (IN CLASS IV ISOFORMS).  
 FT VARSPPLIC 229 241 TTSRIDQAKH -> MNVARAREMLAKF (IN ALPHA  
 FT TYPE ISOFORMS).  
 SQ SEQUENCE 258 AA; 30619 MW; 37BD772A4F28C7E CRC64;  
 Query Match 66.1%; Score 41; DB 1; Length 258;  
 Best Local Similarity 63.6%; Pred. No. 4;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ERAEQRLKQ 11  
 Db 102 ERAEQRLRAE 112  
 TRT3-CHICK  
 ID TRT3-CHICK STANDARD; PRT; 262 AA.  
 AC P12620; P12621; P12619; P12618;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-FEB-1991 (Rel. 17, last sequence update)  
 DT 15-JUN-2002 (Rel. 41, last annotation update)  
 DE Troponin T, fast skeletal muscle isoforms.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 ON NCBI\_TaxId=9031;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8906672; PubMed=3198600;  
 RA Smillie L.B., Golosinska K., Reinach F.C.;  
 RT "Sequences of complete cDNAs encoding four variants of chicken  
 RT skeletal muscle troponin T.";  
 RL J. Biol. Chem. 263:18816-18820(1988).  
 CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF  
 CC TROPONIN. THE THIN FILAMENT REGULATORY COMPLEX WHICH CONFERS  
 CC CALCULUM-SENSITIVITY TO STRIATED MUSCLE ACTINOMYOSIN ATPASE ACTIVITY.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS: TWT-1, TWT-2, TWT-3  
 CC (SHOWN HERE) AND TWT-4; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

CC EMBL: M22156; AAA49102.1; -
DR EMBL: M22156; AAA49103.1; -
DR EMBL: M22155; AAA49101.1; -
DR EMBL: M22154; AAA49100.1; -
DR PIR: A31957; A31957.
DR PIR: B31957; B31957.
DR PIR: C31957; C31957.
DR PIR: D31957; D31957.
DR InterPro: IPR001978; Troponin.
DR Pfam: PF00992; Troponin; 1.
KW Muscle protein; Multigene family; Alternative splicing;
KM Phosphorylation.
FT INIT_MET 0 0
FT MOD_RES 1 1
FT VARSPLIC 11 34
FT VARSPLIC 11 16
FT VARSPLIC 11 16
FT VARSPLIC 38 43
FT VARSPLIC 235 247
SQ SEQUENCE 262 AA; 31010 MW; F6BB67C9ACCET7ED CRC64;

Query Match
Best Local Similarity 66.1%; Score 41; DB 1; Length 262;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEORLRKSO 11
Db 108 ERAEORLRK 118

RESULT 7
TNT3_RABIT
ID TNT3_RABIT STANDARD; PRT; 278 AA.
AC P02641; P19349; P19350;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Troponin T, fast skeletal muscle isoforms.
GN TNT3.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RX MEDLINE-7118575; PubMed-320204;
RA Pearlstone J.R., Johnson P., Carpenter M.R., Smillie I.B.;
RT "Primary structure of rabbit skeletal muscle troponin-T. Sequence
RT determination of the NH2-terminal fragment CB3 and the complete
RT sequence of troponin-T."
RL J. Biol. Chem. 253:983-989(1977).
RN [2]
RX SEQUENCE OF 1-66 FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN-New Zealand white; TISSUE-Skeletal muscle;
RA MEDLINE-93345743; PubMed-8344466;
RX Briggs M.M., Schachar F.;
RT "Origin of fetal troponin T: developmentally regulated splicing of a
RT new exon in the fast troponin T gene."
RL Dev. Biol. 158:503-509(1993).
RN [3]
RX SEQUENCE OF 1-66 FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN-New Zealand white; TISSUE-Skeletal muscle;
RA MEDLINE-90323375; PubMed-2373251;
RX Briggs M.M., McGinnis H.D., Schachar F.;
RT "Transitions from fetal to fast troponin T isoforms are coordinated
RT with changes in tropomyosin and alpha-actinin isoforms in developing
RT rabbit skeletal muscle."
RL Dev. Biol. 140:253-260(1990).

```

```

RN [4]
RX SEQUENCE OF 1-66 FROM N.A., AND ALTERNATIVE SPLICING.
RP STRAIN-New Zealand white; TISSUE-Skeletal muscle;
RA Briggs M.M., Schachar F.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [5]
RX SEQUENCE OF 1-43 AND 56-65.
RC STRAIN-New Zealand white; TISSUE-Skeletal muscle;
RX MEDLINE-89199646; PubMed-2704041;
RA Briggs M.M., Schachar F.;
RT "N-terminal amino acid sequences of three functionally different
RT troponin T isoforms from rabbit fast skeletal muscle."
RL J. Mol. Biol. 206:245-249(1989).
RN [6]
RP PHOSPHORYLATION SITES.
RX MEDLINE-71757104; PubMed-849266;
RA Molr A.J.G., Cole H.A., Perry S.V.;
RT "The phosphorylation sites of troponin T from white skeletal muscle
RT and the effects of interaction with troponin C on their
RT phosphorylation by phosphorylase kinase."
RL Biochem. J. 161:371-382(1977).
CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF
CC TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONFERS
CC CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.
CC -1- ALTERNATIVE PRODUCTS: 9 ISOFORMS: TNT1 (SHOWN HERE), TNT2,
CC TNT2F, TNT2FA, TNT2.5F, TNT3, TNT3F AND TNT4F; ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -1- PTM: SER-168 OR SER-169 AND SER-175 OR SER-176 ARE PHOSPHORYLATED
CC IN VITRO BY PHOSPHORYLASE KINASE.
CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U04970; AAA16023.1; -
DR EMBL: U04971; AAA16024.1; -
DR EMBL: U04972; AAA16025.1; -
DR EMBL: U04973; AAA16026.1; -
DR EMBL: U04974; AAA16027.1; -
DR EMBL: U04975; AAA16028.1; -
DR EMBL: U04976; AAA16029.1; -
DR EMBL: U04977; AAA16030.1; -
DR EMBL: U04978; AAA16031.1; -
DR PIR: A03083; TPBPTS.
DR PIR: S03590; S03590.
DR PIR: S03591; S03591.
DR PIR: S03592; S03592.
DR InterPro: IPR001978; Troponin.
DR Pfam: PF00992; Troponin; 1.
KW Muscle protein; Multigene family; Phosphorylation; Acetylation;
KW Alternative splicing.
FT INIT_MET 0 0
FT MOD_RES 1 1
FT MOD_RES 1 1
FT VARSPLIC 11 16
FT VARSPLIC 22 38
FT VARSPLIC 32 38
FT VARSPLIC 44 55
FT VARSPLIC 19 19
FT CONFLICT 35 36
FT CONFLICT 61 61
SQ SEQUENCE 278 AA; 32903 MW; 3D5A7ESF61E4CA9 CRC64;

```

Query Match 66.1%; Score 41; DB 1; Length 278;  
 Best Local Similarity 63.6%; Pred. No. 4.3;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEOORLSQ 11  
 DB 122 ERAEOORIRAE 132

RESULT 8  
 TRT2\_BOVIN STANDARD; PRT; 284 AA.  
 AC P13789; P13790;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Troponin T, cardiac muscle isoforms (Tnnc).  
 GN TNNT2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_Taxid-9913;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE-88107637; PubMed-3122824;  
 RA Leszyk J., Dumaswala R., Potter J.D., Gusev N.B., Verin A.D.,  
 RA Tobacman L.S., Collins J.H.;  
 RT "Bovine cardiac troponin T: amino acid sequences of the two  
 RT isoforms";  
 RL Biochemistry 26:7035-7042(1987).  
 RN [2]  
 RP PHOSPHORYLATION.  
 RX MEDLINE-90062225; PubMed-2584239;  
 RA Noland T.A. Jr., Raynor R.L., Kuo J.F.;  
 RT "Identification of sites phosphorylated in bovine cardiac troponin I  
 RT and troponin T by protein kinase C and comparative substrate activity  
 RT of synthetic peptides containing the phosphorylation sites";  
 RL J. Biol. Chem. 264:20778-20785(1989).  
 CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF  
 CC TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONFERS  
 CC CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.  
 CC -1- ALTERNATIVE PRODUCTS: 2 Isoforms: 1 (shown here) and 2; are  
 CC produced by alternative splicing.  
 CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.  
 DR PIR: A28008; A28008.  
 DR PIR: B28008; B28008.  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.  
 KW Muscle protein; Alternative splicing; Phosphorylation.  
 FT MOD\_RES 1  
 FT MOD\_RES 190 190  
 FT MOD\_RES 199 199  
 FT MOD\_RES 280 280  
 FT VARSPPLIC 16 21  
 SO SEQUENCE 284 AA; 33782 MW; 08FF3A59A5138530 CRC64;

Query Match 66.1%; Score 41; DB 1; Length 284;  
 Best Local Similarity 63.6%; Pred. No. 4.4;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEOORLSQ 11  
 DB 130 ERAEOORIRAE 140

RESULT 9  
 TRT2\_SHEEP STANDARD; PRT; 286 AA.  
 AC P50751;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Troponin T, cardiac muscle isoform (Tnnc).  
 GN TNNT2.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_Taxid-9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE-Petal heart;  
 RX MEDLINE-91326551; PubMed-1866215;  
 RA McAlliff J.J., Robbins J.;  
 RT "Troponin T expression in normal and pressure-loaded fetal sheep  
 RT heart";  
 RL Pediatr. Res. 29:580-585(1991).  
 CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF  
 CC TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONFERS  
 CC CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM MIDGESTATION TO ADULT LIFE.  
 CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.  
 DR PIR: PF00992; Troponin; 1.  
 DR InterPro: IPR001978; Troponin.  
 KW Muscle protein; Multigene family; Phosphorylation.  
 FT INIT\_MER 0 0  
 FT MOD\_RES 1 1  
 SO SEQUENCE 286 AA; 33633 MW; E59A00F8B3731B1 CRC64;

Query Match 66.1%; Score 41; DB 1; Length 286;  
 Best Local Similarity 63.6%; Pred. No. 4.4;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEOORLSQ 11  
 DB 132 ERAEOORIRAE 142

RESULT 10  
 TRT2\_HUMAN STANDARD; PRT; 297 AA.  
 AC P45379; Q99596; Q99597; O60214;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Troponin T, cardiac muscle isoforms (Tnnc).  
 GN TNNT2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid-9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 6).  
 RP TISSUE-Heart;  
 RC MEDLINE-93345675; PubMed-8344420;  
 RA Mesnard L., Samson F., Espinasse I., Durand J., Neveux J.-Y.,  
 RA Mercadier J.-J.;  
 RT "Molecular cloning and developmental expression of human cardiac  
 RT troponin T";  
 RL FEBS Lett. 328:139-144(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 6).  
 RP TISSUE-Heart muscle;  
 RX MEDLINE-94375053; PubMed-8088824;  
 RA Townsend P.J., Farza H., Macgeoch C., Spurr N.K., Wade R.,  
 RA Gahlman R., Yacoub M.H., Barton P.J.R.;  
 RT "Human cardiac troponin T: Identification of fetal isoforms and  
 RT assignment of the TNNT2 locus to chromosome 1q";  
 RL Genomics 21:311-316(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A. (SPLICED ISOFORMS).  
 RP TISSUE-Fetal heart;  
 RC MEDLINE-96129582; PubMed-8576938;



RA Townsend P.J., Barton P.J.R., Yacoub M.H., Farza H.;  
 RT "Molecular cloning of human cardiac troponin T isoforms: expression in  
 RT developing and failing heart."; J. Mol. Cell. Cardiol. 27:2223-2236(1995).  
 RL J. Mol. Cell. Cardiol. 27:2223-2236(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 6; 7 AND 8).  
 RC TISSUE-Heart;  
 RX MEDLINE-95202803; PubMed-7534662;  
 RA Anderson P.A., Greig A., Mark T.M., Malouf N.N., Oakley A.E.,  
 RA Ungewickler R.M., Allen P.D., Kay B.K.;  
 RT "Molecular basis of human cardiac troponin T isoforms expressed in  
 RT the developing, adult, and failing heart."; J.  
 RL Circ. Res. 76:681-686(1995).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5; 6 AND 10).  
 RC TISSUE-Fetal heart;  
 RX MEDLINE-95202804; PubMed-7895342;  
 RA Mesnard L., Logeart D., Taviaux S., Dirlong S., Mercadier J.-J.,  
 RA Samson F.;  
 RT "Human cardiac troponin T: cloning and expression of new isoforms in  
 RT the normal and failing heart."; J.  
 RL Circ. Res. 76:687-692(1995).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM 6), AND VARIANT FHC IL6-119.  
 RC TISSUE-Heart muscle;  
 RX MEDLINE-98141687; PubMed-9482583;  
 RA Gerull B., Osterziel K.-J., Wilt C., Dietz R., Thierfelder L.;  
 RT "A rapid protocol for cardiac troponin T gene mutation detection in  
 RT familial hypertrophic cardiomyopathy."; J.  
 RL Hum. Mutat. 11:179-182(1998).  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM 6).  
 RA D'Cruz L.G., Oberoi J., Mughal F., Steffensen U., Steffensen M.,  
 RA Kubo T., Mogensen J., McKay G., O'Donoghue A., Pondel M.,  
 RA McKenna W.J., Carter N.D., Baboonian C.;  
 RT "Genomic organization of the human cardiac troponin T gene (TNNT2) and  
 RT characterization of the candidate promoter region."; J.  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE OF 190-228 AND 230-287 FROM N.A.  
 RC TISSUE-Blood;  
 RA Farza H., Townsend P.J.;  
 RT Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP SEQUENCE OF 69-75 AND 176-181.  
 RC TISSUE-Heart;  
 RX MEDLINE-96007936; PubMed-7498159;  
 RA Kovalyov L.I., Shishkin S.S., Efimochkin A.S., Kovalyova M.A.,  
 RA Ershova E.S., Egorov T.A., Musalyanov A.K.;  
 RT "The major protein expression profile and two-dimensional protein  
 RT database of human heart."; J.  
 RL Electrophoresis 16:1160-1169(1995).  
 RN [10]  
 RP VARIANTS FHC ASN-88 AND GLN-101.  
 RX MEDLINE-94265260; PubMed-8205619;  
 RA Thierfelder L., Watkins H., Macrae C., James R., McKenna W.,  
 RA Vosberg H.-P., Seidman J.G., Seidman C.E.;  
 RT "Alpha-tropomyosin and cardiac troponin T mutations cause familial  
 RT hypertrophic cardiomyopathy: a disease of the sarcomere."; J.  
 RL Cell 77:701-712(1994).  
 RN [11]  
 RP VARIANTS FHC.  
 RX MEDLINE-95206332; PubMed-7898523;  
 RA Watkins H., McKenna W.J., Thierfelder L., Suk H.J., Anan R.,  
 RA O'Donoghue A., Spirito P., Matsuura A., Moravec C.S., Seidman J.G.,  
 RA Seidman C.E.;  
 RT "Mutations in the genes for cardiac troponin T and alpha-tropomyosin  
 RT in hypertrophic cardiomyopathy."; J.  
 RL New Engl. J. Med. 332:1058-1064(1995).  
 RN [12]  
 RP VARIANT FHC PRO-287.  
 RA Erdmann J., Wischke S., Kallisch H., Riedel K., Heidenreich M.,  
 RA Fleck E., Regitz-Zagrosek V.;

RT "A novel missense Arg 278 Pro mutation in the troponin T gene  
 RT (TNNT2)."; J. Hum. Mutat. 12:364-364(1998).  
 RL Hum. Mutat. 12:364-364(1998).  
 RN [13]  
 RP VARIANT FHC LEU-103.  
 RX MEDLINE-99457222; PubMed-10525521;  
 RA Varnava A., Baboonian C., Davison F., de Cruz L., Elliott P.M.,  
 RA Davies M.J., McKenna W.J.;  
 RT "A new mutation of the cardiac troponin T gene causing familial  
 RT hypertrophic cardiomyopathy without left ventricular hypertrophy."; J.  
 RL Heart 82:621-624(1999).  
 CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF  
 CC TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONTERS  
 CC CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.  
 CC -1- ALTERNATIVE PRODUCTS: At least 10 isoforms; Isoform 1/TNT1 (shown  
 CC here), 2, 3, 4, 5, 6/TNT3, 7/TNT4, 8/TNT2, 9 and 10; may be  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: Heart. The fetal heart shows a greater  
 CC expression in the atrium than in the ventricle, while the adult  
 CC heart shows a greater expression in the ventricle than in the  
 CC atrium. Isoform 6 predominates in normal adult heart. Isoforms 1,  
 CC 7 and 8 are expressed in fetal heart. Isoform 7 is also expressed  
 CC in failing adult heart.  
 CC -1- DISEASE: DEFECTS IN TNNT2 ARE ONE OF THE CAUSES OF FAMILIAL  
 CC HYPERTROPHIC CARDIOMYOPATHY (FHC) WHICH IS AN AUTOSOMAL DOMINANT  
 CC DISORDER CHARACTERIZED BY INCREASED MYOCARDIAL MASS WITH MYOCYTE  
 CC AND MYOFIBRILLAR DISARRAY. THIS FORM OF FHC IS KNOWN AS CMH2.  
 CC IT IS A DISEASE OF THE SARCOMERE.  
 CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: S64668; AAB27731.1; ALT\_SEQ.  
 DR EMBL: X74819; CA52818.1; -  
 DR EMBL: L40162; AAA67422.1; -  
 DR EMBL: X79855; CA56235.1; -  
 DR EMBL: X79856; CA56236.1; -  
 DR EMBL: X79857; CA56237.1; -  
 DR EMBL: X79858; CA56238.1; -  
 DR EMBL: Y09626; CAA70839.1; -  
 DR EMBL: Y09627; CAA70840.1; -  
 DR EMBL: Y09628; CAA70841.1; -  
 DR EMBL: AF004422; AAC39590.1; -  
 DR EMBL: AF004409; AAC39590.1; JOINED.  
 DR EMBL: AF004410; AAC39590.1; JOINED.  
 DR EMBL: AF004411; AAC39590.1; JOINED.  
 DR EMBL: AF004412; AAC39590.1; JOINED.  
 DR EMBL: AF004413; AAC39590.1; JOINED.  
 DR EMBL: AF004414; AAC39590.1; JOINED.  
 DR EMBL: AF004415; AAC39590.1; JOINED.  
 DR EMBL: AF004416; AAC39590.1; JOINED.  
 DR EMBL: AF004417; AAC39590.1; JOINED.  
 DR EMBL: AF004418; AAC39590.1; JOINED.  
 DR EMBL: AF004419; AAC39590.1; JOINED.  
 DR EMBL: AF004420; AAC39590.1; JOINED.  
 DR EMBL: AF004421; AAC39590.1; JOINED.  
 DR EMBL: AF044273; AAK92231.1; -  
 DR EMBL: AY044273; AAK92231.1; -  
 DR EMBL: S71128; AAB30957.1; -  
 DR EMBL: S71127; AAB30957.1; JOINED.  
 DR Genew; HGNC:11949; TNNT2.  
 DR MIM: 191045; -  
 DR MIM: 115195; -  
 DR InterPro; IPR001978; Troponin.  
 DR Pfam; PF00992; Troponin; 1.  
 DR Muscle protein; Alternative splicing; Multigene family;  
 DR Phosphorylation; Disease mutation; Polymorphism; Cardiomyopathy.  
 FT INIT\_MET 0 0

```

FT MOD_RES 1 1 PHOSPHORYLATION (BY CK2) (BY SIMILARITY).
FT VARSPPLIC 17 21 MISSING (IN ISOFORM 8).
FT VARSPPLIC 17 31 MISSING (IN ISOFORM 7).
FT VARSPPLIC 22 31 MISSING (IN ISOFORM 6).
FT VARSPPLIC 22 22 MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT VARSPPLIC 53 53 MISSING (IN ISOFORM 3 AND ISOFORM 4).
FT VARSPPLIC 98 136 MISSING (IN ISOFORM 9).
FT VARSPPLIC 200 200 MISSING (IN ISOFORM 5).
FT VARSPPLIC 200 202 MISSING (IN ISOFORM 10).
FT VARSPPLIC 88 88 I -> N (IN FHC CMH2).
FT VARIANT 101 101 R -> O (IN FHC CMH2).
FT VARIANT 103 103 /FTID=VAR_007606.
FT VARIANT 119 119 R -> L (IN FHC CMH2).
FT VARIANT 119 119 /FTID=VAR_009194.
FT VARIANT 138 138 F -> I (IN FHC CMH2).
FT VARIANT 138 138 /FTID=VAR_007607.
FT VARIANT 169 169 R -> K.
FT VARIANT 169 169 /FTID=VAR_013021.
FT VARIANT 169 169 MISSING (IN FHC CMH2).
FT VARIANT 169 169 /FTID=VAR_007608.

```

```

Query Match
Best Local Similarity 66.1%; Score 41; DB 1; Length 297;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ERAEODRLKSO 11
DB 142 ERAEODRLKSO 152

```

```

RESULT 11
TRT2_RAT STANDARD; PRT; 298 AA.
AC P50753:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Troponin T, cardiac muscle isoforms (TnTC).
GN TNNI2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89340569; PubMed=2760070;
RA Jin J.-P., Lin J.C.;
RT "Isolation and characterization of cDNA clones encoding embryonic and
RT adult isoforms of rat cardiac troponin T.";
RL J. Biol. Chem. 264:14471-14477(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93059379; PubMed=1433301;
RA Jin J.-P., Huang Q.Q., Yen H.I., Lin J.C.;
RT "Complete nucleotide sequence and structural organization of rat
RT cardiac troponin T gene. A single gene generates embryonic and adult
RT isoforms via developmentally regulated alternative splicing.";
RL J. Mol. Biol. 227:1269-1276(1992).
CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF
CC TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONTERS
CC CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1/Embryonic form (shown here)
CC and 2/adult form; are produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).

```

```

CC -----
CC EMBL: M26051; AAA42296.1; -
CC EMBL: M26052; AAA42297.1; -
CC EMBL: M80829; AAB07676.1; -
CC InterPro: IPR001978; Troponin.
CC Pfam: PF00992; Troponin.1.
CC Muscle protein; Alternative splicing; Multigene family;
CC Phosphorylation.
FT MOD_RES 1 1 BY SIMILARITY.
FT INIT_MET 0 0 PHOSPHORYLATION (BY CK2)
FT VARSPPLIC 17 26 (BY SIMILARITY).
FT CONFLICT 204 204 T -> A (IN REF. 2).
SQ SEQUENCE 298 AA; 35599 MW; 60BE44A8E9DF5E27 CRC64;

```

```

Query Match
Best Local Similarity 66.1%; Score 41; DB 1; Length 298;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ERAEODRLKSO 11
DB 144 ERAEODRLKSO 154

```

```

RESULT 12
TRT2_MOUSE STANDARD; PRT; 300 AA.
AC P50752:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Troponin T, cardiac muscle isoforms (TnTC).
GN TNNI2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Jin J.-P., Wang J., Zhang J.;
RT Tissue-Heart;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF
CC TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONTERS
CC CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; EMBRYONIC/EA (SHOWN HERE) AND
CC A3B: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL: L47553; AAA85349.1; -
CC EMBL: L47600; AAA85352.1; -
CC MGD: MGI:104597; Tnni2.
CC InterPro: IPR001978; Troponin.
CC Pfam: PF00992; Troponin.1.
CC Muscle protein; Alternative splicing; Multigene family;
CC Phosphorylation.
FT MOD_RES 1 1 BY SIMILARITY.
FT INIT_MET 0 0 PHOSPHORYLATION (BY CK2)
FT VARSPPLIC 17 26 (BY SIMILARITY).
FT VARSPPLIC 203 205 MISSING (IN ISOFORM A3B).
SQ SEQUENCE 300 AA; 35694 MW; 3252A8B57E88CE31 CRC64;

```

```

Query Match
Best Local Similarity 66.1%; Score 41; DB 1; Length 300;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

```

RESULT 14
GRPE_STRPY
ID GRPE_STRPY STANDARD; PRI; 190 AA.
AC Q99YC8; P82581;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GRPE protein (HSP-70 cofactor).
GN GRPE OR SPY1761 OR SPYM18.1832.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M18) .
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 186103;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RA Ferrerelli J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Seate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-MGAS8232 / Serotype M18;
RC MEDLINE=21927593; PubMed=11917108;
RX Smoot J.C., Barbhan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Ryan J.C., Barbhan K.D., Van Gompel J.J., Smoot L.M., Zhang Q.,
RA Sylla G.L., Sturdevant D.E., Richlets S.M., Porcella S.F.,
RA Barkins L.D., Bees S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Müsser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
[3]
RN PARTIAL SEQUENCE, AND MASS SPECTROMETRY.
RC STRAIN-JR54 / Serotype M6;
RA Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
RA Vanbogelen R.A.;
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
RT proteins."
RL Submitted (MAY-2000) to the SWISS-PROT data bank.
CC -I- FUNCTION: STIMULATES, JOINTLY WITH DNAI, THE ATPASE ACTIVITY OF
CC DNAA. HELPS TO RELEASE ADP FROM DNAA THUS ALLOWING DNAA TO RECYCLE
CC MORE EFFICIENTLY (BY SIMILARITY).
CC -I- MASS SPECTROMETRY: MM-22054.54; METHOD=Electrospray.
CC -I- SIMILARITY: BELONGS TO THE GRPE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on way
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE006604; AAK34502.1; -
CC EMBL; AE010089; AAL98350.1; -
CC HSSP; P09372; IDKG.
CC InterPro: IPR000740; GRPE.
CC Pfam; PF01025; GRPE; 1.
CC PRINTS; PR00773; GRPEPROTEIN.
CC PROSITE; PS01071; GRPE; 1.
KW Chaperone; Heat shock; Complete proteome.
SQ SEQUENCE 190 AA; 22054 MW; 5636051FC3BAE27E CRC64;
Query Match 62.9%; Score 39; DB 1; Length 190;
Best Local Similarity 61.5%; Pred. No. 6.4;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 ERAEQORLSQDL 13  
 DB 80 ERQOLQRYRSQDL 92

RESULT 15  
 TRT1\_HUMAN STANDARD: PRT: 277 AA.  
 ID P13805:  
 AC 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Troponin T, slow skeletal muscle isoforms (Slow skeletal muscle  
 troponin T).  
 GN TNNT1 OR TNT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88058976; PubMed=2824479;  
 RA Gahlmann R., Trout A.B., Wade R.P., Gunning P., Keddes L.;  
 RT "Alternative splicing generates variants in important functional  
 domains of human slow skeletal troponin T.";  
 RL J. Biol. Chem. 262:16122-16126(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Skeletal muscle;  
 MEDLINE=94183266; PubMed=8135631;  
 RA Samson F., Mesnard L., Mihovilovic M., Potter T.G., Mercadier J.-J.,  
 RA Roses A.D., Gilbert J.R.;  
 RT "A new human slow skeletal troponin T (TnTs) mRNA isoform derived  
 from alternative splicing of a single gene";  
 RL Biochem. Biophys. Res. Commun. 199:841-847(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99208666; PubMed=10191089;  
 RA Barton P.J.R., Cullen M.E., Townsend P.J., Brand N.J., Mullen A.J.,  
 RA Norman D.A.M., Bhavsar P.K., Yacoub M.H.;  
 RT "Close physical linkage of human troponin genes: organization,  
 sequence, and expression of the locus encoding cardiac troponin I and  
 slow skeletal troponin T.";  
 RL Genomics 57:102-109(1999).  
 CC -1- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF  
 TROPONIN. THE THIN FILAMENT REGULATORY COMPLEX WHICH CONFERS  
 CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY.  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M19309; AAA61204.1; -  
 DR EMBL: M19308; AAA61205.1; -  
 DR EMBL: S69208; AAB30272.1; -  
 DR EMBL: S69209; AAB30273.1; -  
 DR EMBL: AJ011712; CAA09751.1; -  
 DR EMBL: AJ011713; CAA09751.1; JOINED.  
 DR EMBL: AJ011712; CAA09752.1; -  
 DR EMBL: AJ011713; CAA09752.1; JOINED.  
 DR PIR: A29783; A29783.  
 DR Genew: HGNC:11948; TNNT1.  
 DR MIM: 191041; -  
 DR InterPro: IPR001978; Troponin.  
 DR Pfam: PF00992; Troponin; 1.

KW Muscle protein; Phosphorylation; Alternative splicing;  
 KW Multigene family.  
 FT INIT\_MET 0  
 FT MOD\_RES 1 1  
 FT VARSPPLIC 24 34 (BY SIMILARITY).  
 FT VARSPPLIC 204 219 MISSING (IN ISOFORM 2).  
 FT CONFLICT 19 19 MISSING (IN ISOFORM 2 AND ISOFORM 3).  
 FT CONFLICT 19 19 E -> D (IN REF. 1).  
 SQ SEQUENCE 277 AA; 32817 MW; 4064FC1F359E663D CRC64;

Query Match 62.9%; Score 39; DB 1; Length 277;  
 Best Local Similarity 63.6%; Pred. No. 9.5;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ERAEQORLSQ 11  
 DB 108 ERAEQORRTE 118

Search completed: March 13, 2003, 18:49:55  
 Job time : 7.85455 secs



```
XX
PS Claim 1; Figure 16; 33pp; Japanese.
XX
CC The cDNA encoding the present sequence, human soluble CD23
CC Isoform D, can be used for the large scale recombinant production
CC of soluble CD23.
CC The specification states that the nucleotide sequences contained
CC in figures 4a-b, 5a-b, 12a-12b and 13a-b are claimed, however this
CC numbering scheme does not relate to the scheme used in the relevant
CC sections of the specification.
XX
SQ Sequence 166 AA;
Query Match 91.9%; Score 57; DB 18; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 RAEQORLKSQDL 13
DB 57 RAEQORLKSQDL 68
RESULT 2
AAW13148
ID AAW13148 standard; Protein; 167 AA.
XX
AC AAW13148;
XX
XX 17-JUN-1997 (first entry)
XX
DE Human soluble CD23 Isoform C.
XX
KW Human; soluble; CD23; complementarity determining;
XX Isoform C; recombinant production.
XX
OS Homo sapiens.
XX
PN JP09028385-A.
XX
PD 04-FEB-1997.
XX
XX 27-DEC-1995; 95JP-0341169.
XX
PR 28-DEC-1994; 94US-0365103.
XX
PA (LYNC/) LYNCH R G.
XX (MATS/) MATSUI M.
XX (NUNE/) NUNEZ R M.
XX (YODO/) YODOI J.
XX
DR WPI: 1997-159094/15.
DR N-PSDB: AAT61961.
XX
PT DNA sequence encoding a soluble Isoform of CD23 - useful for large
PT scale preparation of the protein
XX
PS Claim 1; Figure 15; 33pp; Japanese.
XX
CC The cDNA encoding the present sequence, human soluble CD23
CC Isoform C, can be used for the large scale recombinant production
CC of soluble CD23.
CC The specification states that the nucleotide sequences contained
CC in figures 4a-b, 5a-b, 12a-12b and 13a-b are claimed, however this
CC numbering scheme does not relate to the scheme used in the relevant
CC sections of the specification.
XX
SQ Sequence 167 AA;
Query Match 91.9%; Score 57; DB 18; Length 167;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 RAEQORLKSQDL 13
```

```
DB 58 RAEQORLKSQDL 69
|||||
RESULT 3
AAW13143
ID AAW13143 standard; Protein; 320 AA.
XX
AC AAW13143;
XX
XX 17-JUN-1997 (first entry)
XX
DE Sequence containing human soluble CD23 Isoform B.
XX
KW Human; soluble; CD23; complementarity determining;
XX Isoform B; recombinant production.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Peptide 1..204
XX FT /note="soluble CD23 Isoform B"
XX
PN JP09028385-A.
XX
PD 04-FEB-1997.
XX
XX 27-DEC-1995; 95JP-0341169.
XX
PR 28-DEC-1994; 94US-0365103.
XX
PA (LYNC/) LYNCH R G.
XX (MATS/) MATSUI M.
XX (NUNE/) NUNEZ R M.
XX (YODO/) YODOI J.
XX
DR WPI: 1997-159094/15.
DR N-PSDB: AAT61956.
XX
PT DNA sequence encoding a soluble Isoform of CD23 - useful for large
PT scale preparation of the protein
XX
PS Claim 1; Pages 18-20; 33pp; Japanese.
XX
CC The cDNA encoding the present sequence, which contains human
CC soluble CD23 Isoform B, can be used for the large scale recombinant
CC production of soluble CD23.
CC The specification states that the nucleotide sequences contained
CC in figures 4a-b, 5a-b, 12a-12b and 13a-b are claimed, however this
CC numbering scheme does not relate to the scheme used in the relevant
CC sections of the specification.
XX
SQ Sequence 320 AA;
Query Match 91.9%; Score 57; DB 18; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 RAEQORLKSQDL 13
DB 95 RAEQORLKSQDL 106
RESULT 4
AAP70105
ID AAP70105 standard; Protein; 321 AA.
XX
AC AAP70105;
XX
XX 04-FEB-1991 (first entry)
XX
XX Ige binding factor.
XX
```

KW IGE binding factor; Fc epsilon receptor; allergy; RPMI 8866 cells;  
 XX Homo sapiens.  
 XX EP248211-A.  
 PN  
 XX  
 PD 09-DEC-1987.  
 XX  
 PF 29-APR-1987; 87EP-0106265.  
 XX  
 PR 30-APR-1986; 86JP-0101531.  
 PR 04-SEP-1986; 86JP-0209091.  
 PR 05-SEP-1986; 86JP-0210429.  
 XX  
 PA (KURS ) KURARAY CO. LTD.  
 XX  
 PI Honjo T, Yodot J, Takami M;  
 XX  
 DR WPI: 1987-343202/49.  
 DR N-PSDB; AAN70107.  
 XX  
 PT New IGE binding factor protein - obtd. by culturing cells harbouring a  
 PT class specific Fc receptor on their surface.  
 XX  
 PS Claim 6; page 42-3; 57pp; English.  
 XX  
 CC The protein reacts with anti-Fc epsilon receptor antibody, and is obtd.  
 CC in large quantities. The DNA encoding the protein is obtd. from, eg  
 CC human B cells, human malignant B or T cells, human monocytes or human  
 CC eosinophilic cells. RPMI 8866 cells may also be cultured to produce the  
 CC mRNA. The IGE binding factor and Fc epsilon receptor binds IGE and can  
 CC be used to enhance the prodn. of IGE from IGE producing cells. They can  
 CC be used to treat allergies by eliminating excess blood IGE, and can also  
 SQ Sequence 321 AA;  
 XX  
 QY Query Match 91.9%; Score 57; DB 8; Length 321;  
 Db Best local Similarity 100.0%; Pred. No. 0.1;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 RAEQORLKSODL 13  
 Db 96 RAEQORLKSODL 107  
 XX  
 RESULT 5  
 AAP82073  
 ID AAP82073 standard; protein; 321 AA.  
 XX  
 AC AAP82073;  
 XX  
 DT 17-OCT-1990 (first entry)  
 XX  
 DE Recombinant Fc-epsilon receptor.  
 XX  
 KM Low affinity Fc epsilon receptor; IGE; allergy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..21  
 FT Domain /note="hydrophilic N-terminal"  
 FT Domain 14..20  
 FT Domain /label=stop\_transfer\_seq  
 FT Domain /note="basic cluster involved in  
 FT Domain integration into bilayer"  
 FT Domain 22..47  
 FT Domain /label=transmembrane region  
 FT Domain /note="hydrophobic residues"  
 FT Modified-site 63..63  
 FT /label=glycosylation\_site  
 FT /note="N-linked"  
 FT Cleavage-site 149..150

FT FT /note="target for trypsin-like proteases"  
 FT Protein 1..149  
 FT /label=membrane\_bound  
 FT Protein 150..150  
 FT /label=soluble  
 FT /note="Claim 7"  
 XX  
 PN EP258489-A.  
 XX  
 PD 09-MAR-1988.  
 XX  
 PF 23-SEP-1986; 86EP-0113073.  
 XX  
 PR 21-AUG-1986; 86EP-0111581.  
 XX  
 PA (CELL-) CELLULAR IMMUNOLOGY.  
 XX  
 PI Kishimoto T, Suenmura M, Kikutani H, Barsumian EL;  
 XX  
 DR WPI: 1988-065437/10.  
 DR P-PSDB; AAP82073.  
 XX  
 PT New human low affinity Fc(epsilon)-receptor and parts - useful  
 PT for treating local or systemic allergic reactions and obtd. by  
 PT recombinant DNA methods.  
 XX  
 PS Claim 4; Page 39; 59pp; English.  
 XX  
 CC DNA encoding the sequence or a part of it can be used to make  
 CC recombinant receptor which is useful for treating IGE allergic  
 CC reactions.  
 XX  
 SQ Sequence 321 AA;  
 XX  
 QY Query Match 91.9%; Score 57; DB 9; Length 321;  
 Db Best local Similarity 100.0%; Pred. No. 0.1;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 RAEQORLKSODL 13  
 Db 96 RAEQORLKSODL 107  
 XX  
 RESULT 6  
 AAP81163  
 ID AAP81163 standard; protein; 321 AA.  
 XX  
 AC AAP81163;  
 XX  
 DT 15-OCT-1990 (first entry)  
 XX  
 DE Low affinity Fc-epsilon receptor encoded by gene carried by pGEM4.  
 XX  
 KM Low affinity Fc-epsilon receptor; IGE; allergy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP257114-A.  
 XX  
 PD 02-MAR-1988.  
 XX  
 PF 21-AUG-1986; 86EP-0111581.  
 XX  
 PR 21-AUG-1986; 86EP-0111581.  
 XX  
 PA (KISH/) KISHIMOTO T.  
 XX  
 PI Kishimoto T, Suenmura M, Kikutani H, Barsumian EL;  
 XX  
 DR WPI: 1988-057531/09.  
 DR N-PSDB; AAN81485.  
 XX  
 PT New human lower affinity Fc(epsilon)-receptor - useful for treating

PT local and allergic reactions produced by the expression of IgE.  
 XX  
 PS Claim 9; Page 24; 36pp; English.  
 XX

SQ Sequence 321 AA:

Query Match 91.9%; Score 57; DB 9; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RAEOORLKSQDL 13  
 |||||  
 DB 96 RAEOORLKSQDL 107

RESULT 7  
 AAP81230  
 ID AAP81230 standard; protein; 321 AA.

AC AAP81230;

DT 25-OCT-1990 (first entry)

DE Fc gamma receptor.

KW Fc epsilon receptor; IgE; allergy.

PM Jp63198988-A.

PD 17-AUG-1988.

PF 24-APR-1987; 87JP-0042445.

PR 05-SEP-1986; 86JP-0210429.

PR 24-APR-1987; 87US-0042445.

PA (KURS ) KURARAY KK.

DR WPI: 1988-273895/39

DR N-PSDB; AAN81612.

PT New DNA contg. nucleotide sequence encoding Fc epsilon receptor  
 PT for obtaining IgE-binding substance used to detect and quantify  
 PT IgE responsible for allergies.

PS Disclosure: ge 529; 16pp; Japanese.

CC Fc epsilon receptor binds IgE. It is expressed from cell line RPMI866.  
 CC It is expressed in large amts. and can be used to remove excess IgE from  
 CC blood, and to detect and quantify IgE. This would allow the development  
 CC of therapy and diagnosis of allergy.

SQ Sequence 321 AA:

Query Match 91.9%; Score 57; DB 9; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RAEOORLKSQDL 13  
 |||||  
 DB 96 RAEOORLKSQDL 107

RESULT 8  
 AAP81112

ID AAP81112 standard; peptide; 321 AA.

AC AAP81112;

DT 12-DEC-1990 (first entry)

DE Human IgE binding factor related polypeptide.

XX Human IgE binding factor; IgE-BFs; allergic diseases; B-cells.  
 KW Homo sapiens.  
 XX

OS EP254249-A.

PN 27-JAN-1988.

PD 20-JUL-1987; 87EP-0110458.

PF 22-JUL-1986; 86GB-0017862.

PR 07-NOV-1986; 86GB-0026622.

PA (CIBA ) CIBA GEIGY AG.

PI Hofstetter H, Kilcherr E;

DR WPI: 1988-022917/04.

DR N-PSDB; AAN81437.

PT New IgE binding factor related polypeptide(s) - produced by DNA  
 PT recombinant methods.

PS Claim 1; Page 36; 48pp; English.

CC The peptide is related to the IgE receptors on human B-cells and,  
 CC if without the membrane anchoring sequence, to the IgE-BFs  
 CC of Sarfati et al. Immunology 53, 197, 207, 783 (1984).

CC Amino acids between 1-133 starting from the N-terminal may be deleted.  
 CC This is the membrane-anchoring sequence binding the polypeptide to  
 CC the cytoplasmic membrane of the B-cells.

CC Alternatively amino acids between 110-130 or 250-321 may also be deleted.

CC The peptide has IgE binding activity and is useful for treating  
 CC allergic conditions, e.g. as caused by antigens such as pollens,  
 CC cat danders and house dust mites.

SQ Sequence 321 AA:

Query Match 91.9%; Score 57; DB 9; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RAEOORLKSQDL 13  
 |||||  
 DB 96 RAEOORLKSQDL 107

RESULT 9  
 AAP81172  
 ID AAP81172 standard; protein; 321 AA.

AC AAP81172;

DT 28-MAR-1992 (first entry)

DE Sequence encoded by a gene for the water-soluble part of the  
 DE human low affinity Fc-epsilon-receptor with the amino acids  
 DE 150 to 321, comprising the ECDRI Insert from pFc-epsilon-R-1.

KW Low affinity Fc-epsilon-receptor; systemic IgE-allergic reaction;  
 KW therapy.

OS Homo sapiens.

PN EP259615-A.

PD 16-MAR-1988.

PF 06-AUG-1987; 87EP-0111392.

PR 11-APR-1987; 87EP-0105425.

PR 21-AUG-1986; 86EP-0111581.



PR 23-SEP-1986; 86EP-0113073.  
 PR 05-DEC-1986; 86EP-0116938.  
 PR 06-AUG-1987; 87EP-0111392.  
 XX  
 PA (OSAU ) OSAKA UNIVERSITY.  
 PA (KISH/) KISHIMOTO T.  
 XX  
 PI Kishimoto T, Suemura M, Kikutani H, Barsumian EL;  
 XX  
 DR WPI; 1988-072124/11.  
 DR N-PSDB; AAN81512.  
 XX  
 PT Recombinant human low affinity Fc(epsilon)- receptor - used for  
 PT the treatment of local and systemic IGE-allergic reactions  
 XX  
 PS Claim 3; Table 3, Page 79-81; 118pp; English.  
 XX  
 CC The inventors claim a human low affinity Fc-epsilon-receptor (FCR)  
 CC with an N-terminal cytoplasmic domain, a C-terminal extracellular  
 CC domain and a mol.wt. of about 46kd. Also claimed are recombinant DNA  
 CC contg. the genetic information for the FCR, vectors contg. the DNA,  
 CC host organisms transformed with the vectors, oligonucleotides coding  
 CC for partial amino acid sequences from FCR, and processes for  
 CC preparing FCR. Preferred embodiments of the present invention are  
 CC pFc-epsilon-R-1 (see AAN81516/P81175) and also AAP81172. At least a  
 CC part of the cDNA SO coding for the AAs 1 to 148 of the Fc-epsilon-R  
 CC is replaced by a eucaryotic cDNA signal SO e.g. an interleukin cDNA  
 CC signal SO e.g. by the BSR-2 signal SO (see pBSF-2-L8-AAN81517/P81176).  
 CC A suitable yeast expression vector for is a plasmid contg. the yeast  
 CC ADHI-promoter, a gene coding for the yeast mating factor alpha  
 CC leader peptide (MF alpha leader SO) a multicloning site and the  
 CC yeast ADHI-terminator (see AAN81514/P81173). See also EP-258492.  
 XX  
 SQ Sequence 321 AA:

Query Match 91.9%; Score 57; DB 9; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RAEOQRKLSDDL 13  
 IIIIIIIIIIII  
 DB 96 RAEOQRKLSDDL 107

RESULT 10  
 AAP90120  
 ID AAP90120 standard; protein; 321 AA.  
 XX  
 AC AAP90120;  
 XX  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Human lymphocyte receptor for immunoglobulin.  
 XX  
 KW Human lymphocyte receptor for immunoglobulin; hypersensitivity; allergy;  
 KW asthma; immunoglobulin E; lymphocyte receptor.  
 XX  
 OS Homo sapiens (Human).  
 XX  
 XX EP324879-A.  
 XX  
 PD 26-JUL-1989.  
 XX  
 PF 20-JAN-1988; 88EP-0100814.  
 XX  
 PR 20-JAN-1988; 88EP-0100814.  
 XX  
 PA (OSAU ) OSAKA UNIVERSITY.  
 XX  
 PI Kishimoto T, Suemura M, Kikutani H, Barsumian EL, Schneider F-J;  
 PI Schwendenwein R, Sommergruber W, Swetly P;  
 DR WPI; 1989-214148/30.

DR N-PSDB; AAN90344.  
 XX  
 PT Soluble recombinant Fc-epsilon receptor  
 PT - used for treatment or prophylaxis of local and  
 PT allergic reactions induced by IGE.  
 XX  
 PS Disclosure; fig 1; 23pp; English.  
 XX  
 CC Whole human lymphocyte receptor for immunoglobulin (see corresp.  
 CC AAN90344). Used to produce highly bioactive water-soluble FCR.  
 CC Water-soluble FCR binds IGE, so it is useful for treating  
 CC hypersensitivity, esp. asthma. Amino acid residue 150 is a  
 CC possible site for trypsin-like proteases.  
 XX  
 SQ Sequence 321 AA:

Query Match 91.9%; Score 57; DB 10; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RAEOQRKLSDDL 13  
 IIIIIIIIIIII  
 DB 96 RAEOQRKLSDDL 107

RESULT 11  
 AAP90367  
 ID AAP90367 standard; protein; 321 AA.  
 XX  
 AC AAP90367;  
 XX  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Peptide sequence of mutated Fc epsilon receptor gene.  
 XX  
 KW Fc epsilon receptor; cloned gene; IGE; allergy; asthma.  
 XX  
 PN EP321601-A.  
 XX  
 PD 28-JUN-1989.  
 XX  
 PF 22-DEC-1987; 87EP-0119080.  
 XX  
 PR 22-DEC-1987; 87EP-0119080.  
 XX  
 XX (OSAU ) OSAKA UNIVERSITY.  
 XX  
 PI Kishimoto T, Suemura M, Kikutani H, Barsumian E;  
 XX  
 DR WPI; 1989-186101/26.  
 DR N-PSDB; AAN90134, AAN90135, AAN90136.  
 XX  
 PT Cloned genes coding for soluble IGE receptor - comprising modified  
 PT coding sequence of Fc epsilon receptor gene.  
 XX  
 PS Disclosure; fig 1; 19pp; English.  
 XX  
 CC The known peptide sequence of Fc epsilon receptor gene.  
 CC  
 CC  
 CC  
 SQ Sequence 321 AA:

Query Match 91.9%; Score 57; DB 10; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RAEOQRKLSDDL 13  
 IIIIIIIIIIII  
 DB 96 RAEOQRKLSDDL 107

RESULT 12  
 AAR42053  
 ID AAR42053 standard; protein; 321 AA.

```

XX AC AAK42053;
XX XX 18-FEB-1994 (first entry)
XX DT
XX DE Ige binding factor.
XX KW Ige; binding factor; secretion; expression; signal sequence.
XX OS Synthetic.
XX FH
XX FT Key Location/Qualifiers
XX FT Region 148..321
XX FT /label= Ige-binding_factor
XX PN IL84702-A.
XX PD 08-JUL-1993.
XX XX
XX PF 03-DEC-1987; 87IL-0084702.
XX PR 03-DEC-1987; 87IL-0084702.
XX PA (CIBA ) CIBA GEIGY AG.
XX DR WPI; 1993-275121/35.
XX PT Recombinant DNA molecule - comprises enhancer and promoter unit
XX PT linked to transcriptive DNA segment and DNA segment comprising
XX PT promoter unit linked to DNA sub-segment
XX PS Disclosure; Page 28; 71pp; English.
XX CC The sequence is of Ige-binding factor, secreted from the plasmid
XX CC PSYG-BF. The plasmid PSYG-BF is a derivative of plasmid PSYG-ER
XX CC which has the sequence coding amino acids 1-147 replaced with a new
XX CC coding region (AA053400) allowing secretion of the factor.
XX SQ Sequence 321 AA;

Query Match 91.9%; Score 57; DB 14; Length 321;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RAEQORLKSQDL 13
DB 96 RAEQORLKSQDL 107

RESULT 13
AAW13142
ID AAW13142 standard; Protein; 321 AA.
XX AC
XX AC AAW13142;
XX DT 17-JUN-1997 (first entry)
XX DE
XX DE Sequence containing human soluble CD23 isoform A.
XX KW Human; soluble; CD23; complementarity determining;
XX KW isoform A; recombinant production.
XX OS Homo sapiens.
XX FH
XX FT Key Location/Qualifiers
XX FT Peptide 1..205
XX FT /note= "soluble CD23 isoform A"
XX PN JP09028385-A.
XX PD 04-FEB-1997.
XX PF 27-DEC-1995; 95JP-0341169.

```

```

XX XX 28-DEC-1994; 94US-0365103.
XX PR
XX PA (LYNC/) LYNCH R G.
XX PA (MATSU/) MATSUI M.
XX PA (NUNEZ/) NUNEZ R M.
XX PA (YODO/) YODOI J.
XX DR WPI; 1997-159094/15.
XX DR N-PSDB; AAT61955.
XX PT DNA sequence encoding a soluble isoform of CD23 - useful for large
XX PT scale preparation of the protein
XX PS Claim 1; Pages 16-17; 33pp; Japanese.
XX CC The cDNA encoding the present sequence, which contains human
XX CC soluble CD23 isoform A, can be used for the large scale recombinant
XX CC production of soluble CD23.
XX CC The specification states that the nucleotide sequences contained
XX CC in figures 4a-b, 5a-b, 12a-12b and 13a-b are claimed, however this
XX CC numbering scheme does not relate to the scheme used in the relevant
XX CC sections of the specification.
XX SQ Sequence 321 AA;

Query Match 91.9%; Score 57; DB 18; Length 321;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RAEQORLKSQDL 13
DB 96 RAEQORLKSQDL 107

RESULT 14
AAP82839
ID AAP82839 standard; protein; 336 AA.
XX AC
XX AC AAP82839;
XX DT 16-NOV-1990 (first entry)
XX DE
XX DE Human low affinity Fc(epsilon) receptor.
XX KW Fc(epsilon) receptor; low affinity; Ige-allergic reactions.
XX KW EP258492-A.
XX PN
XX PN 09-MAR-1988.
XX PD
XX PD 05-DEC-1986; 86EP-0116938.
XX PR 05-DEC-1986; 86EP-0116938.
XX PA (CELL-) CELLULAR IMMUNOLOGY.
XX PI Kishimoto T, Suemura M, Kikutani H, Barsumian EI.
XX DR WPI; 1988-065440/10.
XX DR N-PSDB; AAN82252.
XX PT New human low affinity Fc(epsilon)-receptor and parts - useful for
XX PT treating local or systemic Ige-allergic reactions and obd. by
XX PT recombinant DNA methods.
XX PS Disclosure; pp; English.
XX CC This protein is useful in the treatment of local or systemic
XX CC Ige-allergic reactions and is obtd. by recombinant DNA methods. It is
XX CC pref. un- accompanied by associated native glycosylation. Fragments of
XX CC it can also be used. See also AAN82253.

```

SQ Sequence 336 AA;

Job time : 35.5091 secs

Query Match 91.9%; Score 57; DB 9; Length 336;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAEOQRLKSQDL 13

DB 96 RAEOQRLKSQDL 107

RESULT 15

AAP83022

ID AAP83022 standard; protein; 300 AA.

XX AAP83022;

DT 12-DEC-1990 (first entry)

DE Human IGE binding factor related polypeptide from pCL-1.

KW Human IGE binding factor; IGE-BFs; allergic diseases; pCL-1.

XX Homo sapiens.

XX EP54249-A.

PD 27-JAN-1988.

PF 20-JUL-1987; 87EP-0110458.

XX 22-JUL-1986; 86GB-0017862.

PR 07-NOV-1986; 86GB-0026622.

XX (CIBA ) CIBA GEIGY AG.

XX Hofstetter H, Kilcherr E;

DR WPI; 1988-022917/04.

DR N-PSDB; AAN81438.

XX New IGE binding factor related polypeptide(s) - produced by DNA

PT recombinant methods.

XX PS Disclosure; ; 48pp; English.

XX The sequence has the formula of AAP81112 wherein the amino acids 106-127

CC are deleted.

CC The cDNA is obtained by reverse transcription of mRNA isolated from

CC cells expressing IGE-binding activity, esp. human B-cells RPMI 8866.

CC The peptide is related to the IGE receptors on human B-cells and,

CC if without the membrane anchoring sequence, to the IGE-BFs

CC of Sarfati et al. Immunology 53, 197, 207, 783 (1984).

CC Amino acids between 1-133 starting from the N-terminal may be deleted.

CC This is the membrane-anchoring sequence binding the polypeptide to

CC the cytoplasmic membrane of the B-cells.

CC Alternatively amino acids between 110-130 or 250-321 may also be

CC deleted. The peptide has IGE binding activity and is useful for treating

CC allergic conditions, e.g. as caused by antigens such as pollens,

CC cat danders and house dust mites.

XX SO Sequence 300 AA;

QY 2 RAEOQRLKSQDL 13

DB 96 RAEOQRLKSQDL 107

Query Match 85.5%; Score 53; DB 9; Length 300;

Best Local Similarity 91.7%; Pred. No. 0.44;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Search completed: March 13, 2003, 18:49:15

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:17 ; Search time 12.7636 Seconds  
(without alignments)  
97.915 Million cell updates/sec

Title: US-10-050-200-7  
Perfect score: 62  
Sequence: 1 ERAEQGRKSGDL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
1: PIR73:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	91.9	321	1	IGHUR
2	44	71.0	302	1	TPCHTC
3	43	69.4	756	2	ET7590
4	41	66.1	175	2	S48795
5	41	66.1	234	2	B44459
6	41	66.1	249	2	A34327
7	41	66.1	249	2	A44459
8	41	66.1	250	2	S46793
9	41	66.1	251	2	D31957
10	41	66.1	253	2	B34327
11	41	66.1	257	2	B31957
12	41	66.1	258	2	I53021
13	41	66.1	263	2	C31957
14	41	66.1	266	1	TPMBIS
15	41	66.1	272	2	A24824
16	41	66.1	274	2	A31957
17	41	66.1	284	2	A28008
18	41	66.1	289	2	A44781
19	41	66.1	297	2	B44781
20	41	66.1	298	1	TPHUTC
21	41	66.1	299	2	S30443
22	40	64.5	248	2	S10623
23	39	62.9	278	1	TPHUTW
24	39	62.9	309	1	A34198
25	39	62.9	1691	1	A44212
26	38	61.3	146	2	F82418
27	38	61.3	157	2	S13803
28	38	61.3	236	2	E83879
29	38	61.3	266	2	H72665

30	38	61.3	460	2	F84529
31	38	61.3	938	2	A56731
32	38	61.3	1270	2	T09194
33	38	61.3	1534	2	A36734
34	37	59.7	77	2	E69894
35	37	59.7	174	2	A95060
36	37	59.7	182	2	F97928
37	37	59.7	196	2	C82056
38	37	59.7	301	2	T32803
39	37	59.7	312	2	S53969
40	37	59.7	317	2	JC7207
41	37	59.7	383	2	G84597
42	37	59.7	461	2	S50864
43	37	59.7	462	2	A46170
44	37	59.7	495	2	T20754
45	37	59.7	758	2	E75590

## ALIGNMENTS

## RESULT 1

IGHUR  
IE Fc receptor II, low-affinity [validated] - human  
M:Alternate names: Blast-2; CD23; Fc-epsilon-RII; lymphocyte IgE receptor  
N:Contains: Ige Fc receptor II, splice form a; Ige Fc receptor II, splice form a'; Ig  
C:Species: Homo sapiens (man)  
C>Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 15-Sep-2000  
C:Accession: A26067; S03279; S39443; A26164; A26589; A31924; J10132; S29107  
R:Kikutani, H.; Inui, S.; Sato, R.; Barsamian, E.L.; Owaki, H.; Yamasaki, K.; Kaisho, Cell 47, 657-665, 1986  
A>Title: Molecular structure of human lymphocyte receptor for immunoglobulin E.  
A:Reference number: A26067; MUID:87051737; PMID:2877743  
A:Accession: A26067

A:Molecule type: mRNA

A:Residues: 1-321 <KIR>

A:Cross-references: GB:X06049; NID:931316

A:Experimental source: EBV-transformed B lymphoblastoid cells RPMI-8866

R:Suter, U.; Bastos, R.; Hofstetter, H.

Nucleic Acids Res. 15, 7295-7308, 1987

A>Title: Molecular structure of the gene and the 5'-flanking region of the human lymph

A:Reference number: S03279; MUID:88015596; PMID:2958779

A:Accession: S03279

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 157-284 <SUT>

A:Cross-references: GB:X06049; NID:931316

A>Note: all exon sequences were determined but the complete sequence is not shown

R:Matsumi, M.; Nunez, R.; Sachl, Y.; Lynch, R.G.; Yodoi, J.

FEBS Lett. 335, 51-56, 1993

A>Title: Alternative transcripts of the human CD23/Fc-epsilon-RII. A possible novel m

A:Reference number: S39442; MUID:94063078; PMID:8243664

A:Accession: S39442

A:Molecule type: DNA

A:Residues: 1-7, 'D', '47-50 <MAS1>

A:Experimental source: splice form a'

A:Accession: S39443

A:Molecule type: DNA

A:Residues: 'MNPPSD', 47-50 <MAS2>

A:Experimental source: splice form b'

R:Rudin, C.; Hofstetter, H.; Sarfati, M.; Levy, C.A.; Suter, U.; Alaimo, D.; Killeher

EMBO J. 6, 109-114, 1987

A>Title: Cloning and expression of the cDNA coding for a human lymphocyte IgE recepto

A:Reference number: A26164; MUID:87218454; PMID:3034567

A:Accession: A26164

A:Molecule type: mRNA

A:Cross-references: GB:X04772; NID:934002; PIDN:CAA28465.1; PID:934003

A>Note: the codon given for 269-Asn (ACC) is inconsistent with the authors' translati

A>Note: part of this sequence, including the amino end of soluble forms of the protei

R:Ikuta, K.; Takami, M.; Kim, C.W.; Honjo, T.; Miyoshi, T.; Tagaya, Y.; Kawabe, T.; Y

Proc. Natl. Acad. Sci. U.S.A. 84, 819-823, 1987

A>Title: Human lymphocyte Fc receptor for IgE: sequence homology of its cloned cDNA w

A:Reference number: A26589; MUID:87118255; PMID:2949326  
 A:Accession: A26589  
 A:Molecule type: mRNA  
 A:Residues: 1-321 <IRU>  
 A:Cross-references: GB:M15059; NID:g182447; PIDN:AAA52434.1; PID:g182448  
 A:Note: part of this sequence, including the amino end of soluble forms of the protein, R.Yokota, A.; Kikuchi, H.; Tanaka, T.; Sato, R.; Barsmian, E.L.; Suemura, M.; Kishimoto Cell 55, 611-618, 1988  
 A:Title: Two species of human Fc-epsilon1 receptor II (Fc-epsilon1-RII/CD23): tissue-specific  
 A:Reference number: A31924; MUID:89028672; PMID:2972386  
 A:Accession: A31924  
 A:Molecule type: mRNA  
 A:Residues: 'MNP5Q', 8-14 <YOK>  
 A:Cross-references: GB:M23562; NID:g182444  
 A:Experimental source: Splice form IIB  
 R.Letellier, M.; Sarfati, M.; Deleseppe, G.  
 Mol. Immunol. 26, 1105-1112, 1989  
 A:Title: Mechanisms of formation of IGE-binding factors (soluble CD23)-I. Fc epsilon1 R I  
 A:Reference number: JLO132; MUID:90220658; PMID:2534424  
 A:Accession: JLO132  
 A:Molecule type: protein  
 A:Residues: 1-321 <LET>  
 A:Experimental source: lymphoblastoid B cell line  
 R.Rose, K.; Turcatti, G.; Graber, P.; Pochon, S.; Regamey, P.O.; Jansen, K.U.; Magnenat, Biochem. J. 286, 819-824, 1992  
 A:Title: Partial characterization of natural and recombinant human soluble CD23.  
 A:Reference number: S29107; MUID:93038513; PMID:1417742  
 A:Accession: S29107  
 A:Molecule type: protein  
 A:Residues: 152-166;173-179;189-212;230-263;268-306 <ROS>  
 R.Padian, E.A.; Helm, B.A.  
 submitted to the Brookhaven Protein Data Bank, June 1993  
 A:Reference number: A5191; PDB:1HLI  
 A:Contents: annotation; conformation by theoretical model, residues 173-285  
 R.Baforth, J.  
 submitted to the Brookhaven Protein Data Bank, November 1995  
 A:Reference number: A65963; PDB:1KZE  
 A:Contents: annotation; conformation by theoretical model, residues 173-285  
 C:Comment: The sequence of the splice form a is shown.  
 C:Comment: This receptor for the Fc portion of IGE is expressed in various hematopoietic cells.  
 C:Comment: Splice form a is expressed constitutively in B-cells; b is expressed in other cells.  
 C:Comment: Soluble IGE-binding factors are produced by proteolytic cleavage of IGE Fc re  
 C:Genetics:  
 A:Gene: GDB:FCER2; FCE2  
 A:Cross-references: GDB:118888; OMIM:151445  
 A:Map position: 19p13.3-19p13.3  
 A:Introns: 8/1; 46/1; 64/1; 85/1; 106/1; 127/1; 157/1; 207/3; 243/2  
 C:Superfamily: IGE receptor II; C-type lectin homology  
 C:Keywords: alternative splicing; B-cell; glycoprotein; immunoglobulin receptor; macroph  
 F:/MNP5Q/, 47-321/Product: IGE Fc receptor II, splice form b' #status predicted <SFBI>  
 F:/MNP5Q/, 8-321/Product: IGE Fc receptor II, splice form b #status predicted <SRB>  
 F:1-321/Product: IGE Fc receptor II, splice form a #status predicted <SFA>  
 F:1-23/Domain: Intracellular #status predicted <INT>  
 F:1-7, 'D', 47-321/Product: IGE Fc receptor II, splice form a' #status predicted <SFAI>  
 F:14-20/Region: Stop-transfer sequence  
 F:22-45/Domain: transmembrane #status predicted <TM>  
 F:46-321/Domain: extracellular #status predicted <EXT>  
 F:64-84/Region: 21-residue repeat  
 F:81-321/Product: soluble IGE-binding factor (37K) #status predicted <IGE>  
 F:85-105/Region: 21-residue repeat  
 F:102-321/Product: soluble IGE-binding factor (33K) #status predicted <IGB>  
 F:106-126/Region: 21-residue repeat  
 F:125-321/Product: soluble IGE-binding factor (29K) #status predicted <IGI>  
 F:148-321/Product: soluble IGE-binding factor (25-27K), long form #status experimental  
 F:150-321/Product: soluble IGE-binding factor (25-27K), short form #status experimental  
 F:163-282/Domain: C-type lectin homology <LCH>  
 F:63/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:149-148/Cleavage site: lys-Met (unidentified proteinase) #status experimental  
 F:191-282,259-273/Disulfide bonds: #status experimental

Query Match 91.9%; Score 57; DB 1; Length 321;

Best Local Similarity 100.0%; Pred. No. 0.016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 RAEOORLKSODL 13  
 Db 96 RAEOORLKSODL 107

## RESULT 2

TPCHRC

troponin T, cardiac muscle, embryonic splice form - chicken  
 N:Contents: cardiac muscle troponin T, adult splice form  
 C:Species: Gallus gallus (chicken)  
 C:Date: 28-Aug-1985 #sequence\_revision 12-Apr-1996 #text\_change 16-Jul-1999  
 C:Accession: A25373; A03086

J.Cooper, T.A.; Ordahl, C.P.  
 J. Biol. Chem. 260, 11140-11148, 1985  
 A:Title: A single cardiac troponin T gene generates embryonic and adult isoforms via  
 A:Reference number: A25373; MUID:85289327; PMID:2993302  
 A:Accession: A25373

A:Molecule type: mRNA  
 A:Residues: 1-302 <COO>  
 A:Cross-references: GB:M10013; NID:g212783; PIDN:AAA49099.1; PID:g212784  
 R.Cooper, T.A.; Ordahl, C.P.  
 Science 226, 979-982, 1984

A:Title: A single troponin T gene regulated by different programs in cardiac and skel  
 A:Reference number: A03086; MUID:85065747; PMID:6095446  
 A:Accession: A03086

A:Molecule type: mRNA  
 A:Residues: 68-302 <CO2>  
 A:Cross-references: GB:K02263; NID:g212781; PIDN:AAA49098.1; PID:g212782  
 C:Comment: This protein, found in adult cardiac muscle and transiently in embryonic s  
 C:Superfamily: troponin T

C:Keywords: alternative splicing; cardiac muscle; differentiation; heart; muscle; pho  
 F:1-302/Product: troponin T, cardiac muscle, embryonic splice form #status predicted  
 F:1-22,33-302/Product: troponin T, cardiac muscle, adult splice form #status predicted

## Query Match

Best Local Similarity 71.0%; Score 44; DB 1; Length 302;  
 Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ERAEOORLKSQ 11  
 Db 147 ERAEOORLRS 157

## RESULT 3

E75590  
 methyl-accepting chemotaxis-related protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: E75590  
 R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
 M.; Shen, H.O.; Vamshayan, J.D.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: E75590

A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-756 <WHI>  
 A:Cross-references: GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF12432.1; PID:g646  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DRA0353

## Query Match

Best Local Similarity 69.4%; Score 43; DB 2; Length 756;  
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;



Query Match 66.1%; Score 41; DB 2; Length 250;  
 Best Local Similarity 63.6%; Pred. No. 8.2;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEOORLKSQ 11  
 DB 128 ERAEOORIRAE 138

# RESULT 9

D31957

tropoin T, skeletal muscle, isoform 4 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 13-Aug-1999

C:Accession: D31957

R:Smillie, L.B.; Golosinska, K.; Reinach, F.C.

J. Biol. Chem. 263, 18816-18820, 1988

A:Title: Sequences of complete cDNAs encoding four variants of chicken skeletal muscle

A:Reference number: A92698; MUID:8906672; PMID:3198600

A:Accession: D31957

A:Molecule type: mRNA

A:Residues: 1-251 <SMT>

A:Cross-references: GB:M2156; GB:J04198; NID:g212791; PIDN:AAA49103.1; PID:g212792

C:Superfamily: tropoin T

C:Keywords: alternative splicing; muscle; phosphoprotein; skeletal muscle

Query Match

Best Local Similarity 66.1%; Score 41; DB 2; Length 251;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEOORLKSQ 11  
 DB 97 ERAEOORIRAE 107

# RESULT 10

B34327

tropoin T, fast skeletal muscle, adult alpha (clone 605) - Japanese quail

C:Species: Coturnix coturnix japonica (Japanese quail)

C:Date: 22-Jun-1990 #sequence\_revision 22-Jun-1990 #text\_change 13-Aug-1999

C:Accession: B34327

R:Butcher, E.A.; Charles de la Brousse, F.; Emerson Jr., C.P.

J. Biol. Chem. 264, 12482-12491, 1989

A:Title: Developmental and muscle-specific regulation of avian fast skeletal tropoin T

A:Reference number: A34327; MUID:8930680; PMID:2745456

A:Accession: B34327

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-253 <BUC>

A:Cross-references: GB:M2600; NID:g213629; PIDN:AAA49506.1; PID:g213630; GB:J05006

C:Superfamily: tropoin T

C:Keywords: skeletal muscle

Query Match

Best Local Similarity 66.1%; Score 41; DB 2; Length 253;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEOORLKSQ 11  
 DB 99 ERAEOORIRAE 109

# RESULT 11

B31957

tropoin T, skeletal muscle, isoform 2 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 13-Aug-1999

C:Accession: B31957

R:Smillie, L.B.; Golosinska, K.; Reinach, F.C.

J. Biol. Chem. 263, 18816-18820, 1988

A:Title: Sequences of complete cDNAs encoding four variants of chicken skeletal muscle

A:Reference number: A92698; MUID:8906672; PMID:3198600

A:Accession: B31957

A:Molecule type: mRNA  
 A:Residues: 1-257 <SMT>  
 A:Cross-references: GB:M2155; GB:J04198; NID:g212787; PIDN:AAA49101.1; PID:g212788  
 C:Superfamily: tropoin T  
 C:Keywords: alternative splicing; muscle; phosphoprotein; skeletal muscle

Query Match 66.1%; Score 41; DB 2; Length 257;  
 Best Local Similarity 63.6%; Pred. No. 8.4;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEOORLKSQ 11  
 DB 103 ERAEOORIRAE 113

# RESULT 12

I53021

tropoin T - human

C:Species: Homo sapiens (man)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 13-Aug-1999

C:Accession: I53021

R:Wu, Q.L.; Jha, P.K.; Raychowdhury, M.K.; Du, Y.; Leavis, P.C.; Sarkar, S.

DNA Cell Biol. 13, 217-233, 1994

A:Title: Isolation and characterization of human fast skeletal beta tropoin T cDNA:

A:Reference number: I53021; MUID:94226695; PMID:8172653

A:Accession: I53021

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-258 <RES>

A:Cross-references: GB:M21984; NID:g508854; PIDN:AAA36777.1; PID:g508855

C:Genetics:

A:Gene: GDB:TNNT3

A:Cross-references: GDB:350962; OMIM:600692

A:Map position: 11p15.5-11p15.5

C:Superfamily: tropoin T

Query Match

Best Local Similarity 66.1%; Score 41; DB 2; Length 258;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEOORLKSQ 11  
 DB 102 ERAEOORIRAE 112

# RESULT 13

C31957

tropoin T, skeletal muscle, isoform 3 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 13-Aug-1999

C:Accession: C31957

R:Smillie, L.B.; Golosinska, K.; Reinach, F.C.

J. Biol. Chem. 263, 18816-18820, 1988

A:Title: Sequences of complete cDNAs encoding four variants of chicken skeletal muscle

A:Reference number: A92698; MUID:8906672; PMID:3198600

A:Accession: C31957

A:Molecule type: mRNA

A:Residues: 1-263 <SMT>

A:Cross-references: GB:M2156; GB:J04198; NID:g212789; PIDN:AAA49102.1; PID:g212790

C:Superfamily: tropoin T

C:Keywords: alternative splicing; muscle; phosphoprotein; skeletal muscle

Query Match

Best Local Similarity 66.1%; Score 41; DB 2; Length 263;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEOORLKSQ 11  
 DB 109 ERAEOORIRAE 119

# RESULT 14

TPRBYS



troponin T, fast skeletal muscle - rabbit  
 N:Contin: troponin T, fast skeletal muscle splice form 1; troponin T, fast skeletal mu  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C>Date: 24-Apr-1984 #sequence\_revision 17-May-1996 #text\_change 22-Jun-1999  
 C:Accession: A03083; S03590; S03591; S03592; I46515; I46516  
 R:Pearlstone, J.R.; Johnson, P.; Carpenter, M.R.; Smillie, L.B.  
 J. Biol. Chem. 252, 983-989, 1977  
 A:Title: Primary structure of rabbit skeletal muscle troponin-T. Sequence determination  
 A:Reference number: A92220; MUID:77118575; PMID:320204  
 A:Accession: A03083  
 A:Molecule type: protein  
 A:Residues: 1-10,17-18,'E',20-48,50-266 <PE2>  
 A:Note: this is the final paper in a series  
 R:Briggs, M.M.; Schachar, F.  
 J. Mol. Biol. 206, 245-249, 1989  
 A:Title: N-terminal amino acid sequences of three functionally different troponin T iso  
 A:Reference number: S03590; MUID:89199646; PMID:2704041  
 A:Accession: S03590  
 A:Molecule type: protein  
 A:Residues: 1-10,17-53 <BR1>  
 A:Note: splice form 2  
 A:Accession: S03591  
 A:Molecule type: protein  
 A:Residues: 1-53 <BR2>  
 A:Note: splice form 1  
 A:Accession: S03592  
 A:Molecule type: protein  
 A:Residues: 1-21,39-53 <BR3>  
 R:Molt, A.J.G.; Cole, H.A.; Perry, S.V.  
 Biochem. J. 161, 371-382, 1977  
 A:Title: The phosphorylation sites of troponin T from white skeletal muscle and the effe  
 A:Reference number: A90295; MUID:77157104; PMID:849266  
 A:Contents: annotation; phosphorylation sites  
 A:Note: phosphorylation under in vivo conditions occurred at positions that may not be a  
 R:Putney, S.D.; Herlihy, W.C.; Schimmel, P.  
 Nature 302, 718-721, 1983  
 A:Title: A new troponin T and cDNA clones for 13 different muscle proteins, found by sh  
 A:Reference number: I46471; MUID:83167564; PMID:6687628  
 A:Accession: I46515  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 56-136 <PUT>  
 A:Cross-references: EMBL:V00899; NID:g1740; PIDN:CAA24264.1; PID:g929768  
 A:Accession: I46516  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 240-266 <PU2>  
 A:Cross-references: EMBL:V00900; NID:g1742; PIDN:CAA24265.1; PID:g833794  
 C:Complex: troponin is a heterotrimer with one molecule each of troponin C (calcium bind  
 C:Function:  
 A:Description: binds the troponin complex to tropomyosin; with tropomyosin mediates cont  
 A:Pathway: muscle contraction  
 C:Superfamily: troponin T  
 C:Keywords: acetylated amino end; alternative splicing; differentiation; muscle contract  
 F:1-266/Product: troponin T, fast skeletal muscle splice form 1 #status experimental <M  
 F:1-21,39-266/Product: troponin T, fast skeletal muscle splice form 3 #status experiment  
 F:1-10,17-266/Product: troponin T, fast skeletal muscle splice form 2 #status experiment  
 F:1/Modified site: acetylated amino end (Ser) #status experimental  
 F:1/Binding site: phosphate (Ser) (covalent) (by troponin T kinase) #status experimental

Query Match 66.1% Score 41; DB 1; Length 266;  
 Best Local Similarity 63.6% Pred. NO. 8.7;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLKSQ 11  
 |||||  
 Db 110 ERAEQORLRAE 120

RESULT 15  
 A24824  
 troponin T, fast skeletal muscle - rat  
 C:Species: Rattus norvegicus (Norway rat)

C>Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Feb-1997  
 C:Accession: A24824  
 R:Breitbart, R.E.; Nadal-Ginard, B.  
 J. Mol. Biol. 188, 313-324, 1986  
 A:Title: Complete nucleotide sequence of the fast skeletal troponin T gene. Alternat  
 A:Reference number: A24824; MUID:86281691; PMID:3735424  
 A:Accession: A24824  
 A:Molecule type: DNA  
 A:Residues: 1-272 <BRE>  
 C:Genetics:  
 A:Introns: 6/2; 11/1 17/1; 23/1; 28/1; 32/1; 37/1; 43/2; 58/3; 97/3; 123/3; 161/3; 1  
 C:Superfamily: troponin T  
 C:Keywords: alternative splicing; muscle; phosphoprotein; skeletal muscle

Query Match 66.1% Score 41; DB 2; Length 272;  
 Best Local Similarity 63.6% Pred. NO. 9;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERAEQORLKSQ 11  
 |||||  
 Db 103 ERAEQORLRAE 113

Search completed: March 13, 2003, 18:53:00  
 Job time : 13.7636 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:52 ; Search time 11.5818 Seconds  
(without alignments)  
33.026 Million cell updates/sec

Title: US-10-050-200-7

Perfect score: 62  
Sequence: 1 ERAEQRLKSQDL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database :

Issued Patents\_AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/iaa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	91.9	166	1	US-08-365-103B-14
2	57	91.9	167	1	US-08-365-103B-12
3	57	91.9	320	1	US-08-365-103B-10
4	57	91.9	321	1	US-08-365-103B-8
5	57	91.9	321	1	US-08-365-103B-8
6	41	66.1	258	2	US-08-602-941-3
7	41	66.1	258	2	US-08-961-264-3
8	41	66.1	258	4	US-09-442-099A-3
9	41	66.1	258	4	US-08-961-858-6
10	41	66.1	258	3	US-09-089-593-6
11	40	64.5	208	4	US-08-850-925-4
12	40	64.5	208	4	US-09-535-521-20
13	40	64.5	292	4	US-09-535-521-2
14	37	59.7	174	4	US-08-365-103B-6
15	37	59.7	174	4	US-08-986-967-2
16	37	59.7	174	4	US-08-858-207A-293
17	37	59.7	317	2	US-08-973-275-1
18	37	59.7	310	1	US-08-249-112-3
19	36	58.1	510	5	PCT-US95-06556-3
20	36	58.1	300	1	US-08-365-103B-4
21	36	58.1	327	1	US-08-365-103B-6
22	36	58.1	724	4	US-08-365-103B-2
23	35	56.5	162	4	US-09-562-737-27
24	35	56.5	162	4	US-09-370-838-93
25	35	56.5	276	1	US-07-612-674-5
26	35	56.5	279	1	US-07-612-674-8
27	34	54.8	329	4	US-09-443-184-56
					Sequence 7, Appl

28	34	54.8	362	2	US-09-080-897-6	Sequence 6, Appl
29	34	54.8	362	4	US-09-323-735-6	Sequence 6, Appl
30	34	54.8	431	3	US-08-478-507-2	Sequence 2, Appl
31	34	54.8	431	4	US-09-128-275A-2	Sequence 2, Appl
32	34	54.8	431	4	US-09-553-427-2	Sequence 2, Appl
33	34	54.8	724	4	US-09-562-737-21	Sequence 21, Appl
34	34	54.8	724	4	US-09-562-737-24	Sequence 24, Appl
35	34	54.8	724	4	US-09-562-737-29	Sequence 29, Appl
36	34	54.8	840	4	US-08-974-549A-190	Sequence 190, App
37	34	54.8	872	3	US-08-851-843A-8	Sequence 8, Appl
38	34	54.8	872	3	US-08-851-843A-54	Sequence 54, Appl
39	34	54.8	872	4	US-08-974-549A-221	Sequence 221, App
40	34	54.8	872	4	US-08-854-050-8	Sequence 8, Appl
41	34	54.8	872	4	US-08-854-050-54	Sequence 54, Appl
42	34	54.8	872	4	US-09-430-323-8	Sequence 8, Appl
43	34	54.8	872	4	US-09-430-323-54	Sequence 54, Appl
44	34	54.8	1248	2	US-09-080-897-2	Sequence 2, Appl
45	34	54.8	1248	4	US-09-323-735-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-08-365-103B-14  
Sequence 14, Application US/08365103B  
Patent No. 5766943  
GENERAL INFORMATION:  
APPLICANT: Lynch, Richard G  
APPLICANT: Nunez, Raphael D.  
APPLICANT: Yodot, Jungi  
TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23  
TITLE OF INVENTION: and Methods of Use for Same  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease  
STREET: 801 Grand Ave. Suite 3200  
CITY: Des Moines  
STATE: Iowa  
COUNTRY: United States  
ZIP: 50309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/365,103B  
FILING DATE: 28-DEC-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Nebel, Heidi S.  
REGISTRATION NUMBER: 37,719  
REFERENCE/DOCKET NUMBER: Uif NS-24  
TELEPHONE: (515) 288-3667  
TELEFAX: (515) 288-1338  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 166 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-365-103B-14

Query Match 91.9%; Score 57; DB 1; Length 166;  
Best Local Similarity 100.0%; Pred. No. 0.015; 0; Indels 0; Gaps 0;  
Matches 12; Conservative 0; Mismatches 0;

QY 2 RAEOQLKSQDL 13  
|||||||  
DB 57 RAEOQLKSQDL 68

```
RESULT 2
US-08-365-103B-12
; Sequence 12, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G
; APPLICANT: Nunez, Raphael D.
; TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave. Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365.103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Ulf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-365-103B-12
;
Query Match          91.9%; Score 57; DB 1; Length 167;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RAEQRLKSQDL 13
      |||||||
Db      58 RAEQRLKSQDL 69

RESULT 3
US-08-365-103B-10
; Sequence 10, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G
; APPLICANT: Nunez, Raphael D.
; TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave. Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365.103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Ulf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-365-103B-8
;
Query Match          91.9%; Score 57; DB 1; Length 321;
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365.103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Ulf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-365-103B-10
;
Query Match          91.9%; Score 57; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RAEQRLKSQDL 13
      |||||||
Db      95 RAEQRLKSQDL 106

RESULT 4
US-08-365-103B-8
; Sequence 8, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G
; APPLICANT: Nunez, Raphael D.
; TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave. Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365.103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Ulf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-365-103B-8
;
Query Match          91.9%; Score 57; DB 1; Length 321;
```

Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ERAEQRLKSQDL 13  
|||||  
Db 96 ERAEQRLKSQDL 107

## RESULT 5

US-08-602-941-3  
; Sequence 3, Application US/08602941  
; Patent No. 5837680  
; GENERAL INFORMATION:  
; APPLICANT: Moses, Marsha A.  
; APPLICANT: Langer, Robert S.  
; APPLICANT: Wiederschain, Dimitri G.  
; APPLICANT: Wu, Immin  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
; TITLE OF INVENTION: COMPRISING TROPONIN SUBUNITS, FRAGMENTS AND ANALOGS  
; TITLE OF INVENTION: THEREOF AND METHODS OF THEIR USE TO INHIBIT ANGIOGENESIS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: IBM Compatible  
; SOFTWARE: FASTSEQ Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/602,941  
; FILING DATE: 16-FEB-1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Poissant, Brian M.  
; REGISTRATION/DOCKET NUMBER: 28,452  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864/9741  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 258 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 5837680e  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..258  
; OTHER INFORMATION: /label= Human Fast Skeletal Beta  
; OTHER INFORMATION: Troponin T  
US-08-602-941-3

Query Match 66.1%; Score 41; DB 2; Length 258;  
Best Local Similarity 63.6%; Pred. No. 10;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERAEQRLKSQ 11  
|||||  
Db 102 ERAEQRLKSQ 112

RESULT 6  
US-08-961-264-3  
; Sequence 3, Application US/08961264  
; Patent No. 6025331  
; GENERAL INFORMATION:

; APPLICANT: Moses, Marsha A.  
; APPLICANT: Langer, Robert S.  
; APPLICANT: Wiederschain, Dimitri G.  
; APPLICANT: Wu, Immin  
; APPLICANT: Sytkowski, Arthur  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
; TITLE OF INVENTION: COMPRISING TROPONIN SUBUNITS, FRAGMENTS AND ANALOGS  
; TITLE OF INVENTION: THEREOF AND METHODS OF THEIR USE TO INHIBIT ANGIOGENESIS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: IBM Compatible  
; SOFTWARE: FASTSEQ Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,264  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/602,941  
; FILING DATE: 16-FEB-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Poissant, Brian M.  
; REGISTRATION/DOCKET NUMBER: 28,452  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864/9741  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 258 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6025331e  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..258  
; OTHER INFORMATION: /label= Human Fast Skeletal Beta  
; OTHER INFORMATION: Troponin T  
US-08-961-264-3

Query Match 66.1%; Score 41; DB 3; Length 258;  
Best Local Similarity 63.6%; Pred. No. 10;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERAEQRLKSQ 11  
|||||  
Db 102 ERAEQRLKSQ 112

RESULT 7  
US-09-442-099A-3  
; Sequence 3, Application US/09442099A  
; Patent No. 6465431  
; GENERAL INFORMATION:  
; APPLICANT: Thorn, R.  
; APPLICANT: Langer, M.  
; APPLICANT: Moses, M.  
; APPLICANT: Wiederschain, D.  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING TROPONIN SUBUNITS,  
; TITLE OF INVENTION: FRAGMENTS AND HOMOLOGS THEREOF AND METHODS OF THEIR USE TO  
; TITLE OF INVENTION: INHIBIT ANGIOGENESIS  
; FILE REFERENCE: 8657-028  
; CURRENT APPLICATION NUMBER: US/09/442,099A  
; CURRENT FILING DATE: 1997-11-17

;; PRIOR APPLICATION NUMBER: 09/268,274  
;; PRIOR FILING DATE: 1999-03-15  
;; PRIOR APPLICATION NUMBER: 08/961,264  
;; PRIOR FILING DATE: 1997-10-30  
;; PRIOR APPLICATION NUMBER: 08/602,941  
;; PRIOR FILING DATE: 1996-02-16  
;; NUMBER OF SEQ ID NOS: 20  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO: 3  
;; LENGTH: 258  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-442-099A-3

Query Match 66.1%; Score 41; DB 4; Length 258;  
Best Local Similarity 63.6%; Pred. No. 10;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERAEQORLSQ 11  
|||||  
DB 102 ERAEQORIRAE 112

RESULT 8  
US-08-961-858-6  
; Sequence 6, Application US/08961858  
; Patent No. 5834210  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Shigu  
; APPLICANT: Shi, Qimwei  
; TITLE OF INVENTION: STABLE TROPONIN SUBUNITS AND COMPLEXES  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,858  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1112-1-044 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 288 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-08-961-858-6

Query Match 66.1%; Score 41; DB 2; Length 288;  
Best Local Similarity 63.6%; Pred. No. 11;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERAEQORLSQ 11  
|||||  
DB 133 ERAEQORIRAE 143

RESULT 9  
US-09-089-593-6  
; Sequence 6, Application US/09089593  
; Patent No. 6060278  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Shigu  
; APPLICANT: Shi, Qimwei  
; TITLE OF INVENTION: STABLE TROPONIN SUBUNITS AND COMPLEXES  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/089,593  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,858  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1112-1-044 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 288 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-09-089-593-6

Query Match 66.1%; Score 41; DB 3; Length 288;  
Best Local Similarity 63.6%; Pred. No. 11;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERAEQORLSQ 11  
|||||  
DB 133 ERAEQORIRAE 143

RESULT 10  
US-08-950-925-4  
; Sequence 4, Application US/08950925  
; Patent No. 6072040  
; GENERAL INFORMATION:  
; APPLICANT: Dave, Kirtl I.  
; APPLICANT: Botyanski, Janos  
; APPLICANT: Sinter, Eva  
; TITLE OF INVENTION: Stabilized Conjugates of Uncomplexed  
; TITLE OF INVENTION: Subunits of Multimeric Proteins  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA

COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/950,925  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Akford, Laurie A  
REGISTRATION NUMBER: 35,053  
REFERENCE/DOCKET NUMBER: 32260-20004.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-813-5600  
TELEFAX: 650-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 288 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-950-925-4

Query Match 66.1%; Score 41; DB 3; Length 288;  
Best Local Similarity 63.6%; Pred. No. 11;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERAEQRLKSD 11  
Db 133 ERAEQRLKSD 143

RESULT 11  
US-09-535-521-20  
Sequence 20, Application US/09535521  
Patent No. 6410714  
GENERAL INFORMATION:  
APPLICANT: Weber, Eric R.  
TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF  
FILE REFERENCE: AL-5  
CURRENT APPLICATION NUMBER: US/09/535,521  
CURRENT FILING DATE: 2000-03-24  
EARLIER APPLICATION NUMBER: 60/125,913  
EARLIER FILING DATE: 1999-03-24  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 20  
LENGTH: 208  
TYPE: PRT  
ORGANISM: Canis familiaris  
US-09-535-521-20

Query Match 64.5%; Score 40; DB 4; Length 208;  
Best Local Similarity 63.6%; Pred. No. 12;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAEOQRLKSD 12  
Db 16 QAEOQRLKSD 26

RESULT 12

US-09-535-521-2  
Sequence 2, Application US/09535521  
Patent No. 6410714  
GENERAL INFORMATION:  
APPLICANT: Weber, Eric R.  
TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF  
FILE REFERENCE: AL-5  
CURRENT APPLICATION NUMBER: US/09/535,521  
CURRENT FILING DATE: 2000-03-24  
EARLIER APPLICATION NUMBER: 60/125,913  
EARLIER FILING DATE: 1999-03-24  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 292  
TYPE: PRT  
ORGANISM: Canis familiaris  
US-09-535-521-2

Query Match 64.5%; Score 40; DB 4; Length 292;  
Best Local Similarity 63.6%; Pred. No. 17;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAEOQRLKSD 12  
Db 100 QAEOQRLKSD 110

RESULT 13  
US-09-535-521-5  
Sequence 5, Application US/09535521  
Patent No. 6410714  
GENERAL INFORMATION:  
APPLICANT: Weber, Eric R.  
TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF  
FILE REFERENCE: AL-5  
CURRENT APPLICATION NUMBER: US/09/535,521  
CURRENT FILING DATE: 2000-03-24  
EARLIER APPLICATION NUMBER: 60/125,913  
EARLIER FILING DATE: 1999-03-24  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 292  
TYPE: PRT  
ORGANISM: Canis familiaris  
US-09-535-521-5

Query Match 64.5%; Score 40; DB 4; Length 292;  
Best Local Similarity 63.6%; Pred. No. 17;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAEOQRLKSD 12  
Db 100 QAEOQRLKSD 110

RESULT 14  
US-08-986-967-2  
Sequence 2, Application US/08986967  
Patent No. H002023  
GENERAL INFORMATION:  
APPLICANT: Hoskins, Jo A.  
APPLICANT: Jaskunas Jr., Stanley R.  
APPLICANT: Rockey, Pamela K.  
APPLICANT: Treadway, Paul J.  
TITLE OF INVENTION: Streptococcus Pneumoniae Gene Sequence  
NUMBER OF SEQUENCES: 4

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Eli Lilly and Company  
;; STREET: Lilly Corporate Center  
;; CITY: Indianapolis  
;; STATE: Indiana  
;; COUNTRY: U.S.  
;; ZIP: 46285  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent Release #1.0, Version #1.30  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/966,967  
;;  
;; FILING DATE:  
;; CLASSIFICATION: 536  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Webster, Thomas D.  
;;  
;; REGISTRATION NUMBER: 39,872  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 317-276-3334  
;; TELEFAX: 317-276-2763  
;;  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 174 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;;  
US-08-966-967-2

Query Match 59.7%; Score 37; DB 1; Length 174;  
Best Local Similarity 61.5%; Pred. No. 31;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 ERAEQRLKSODL 13  
DB 63 ERONLQRYRSODL 75

RESULT 15  
US-08-858-207A-293  
; Sequence 293, Application US/08858207A  
; Patent No. 6348328  
; GENERAL INFORMATION:  
; APPLICANT: Black, Michael  
; APPLICANT: Hodgson, John  
; APPLICANT: Knowles, David  
; APPLICANT: Nichols, Richard  
; APPLICANT: Stodola, Robert  
; TITLE OF INVENTION: No. 6348328e1 Compounds  
; NUMBER OF SEQUENCES: 552  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/858,207A  
; FILING DATE: 09-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/017670  
; FILING DATE: 14-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gimm, Edward R

;; REGISTRATION NUMBER: 38,891  
;; REFERENCE/DOCKET NUMBER: P50475  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 610-270-4478  
;; TELEFAX: 610-270-5090  
;;  
;; TELEX:  
;;  
;; INFORMATION FOR SEQ ID NO: 293:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 174 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: NO. 6348328e  
;;  
US-08-858-207A-293

Query Match 59.7%; Score 37; DB 4; Length 174;  
Best Local Similarity 61.5%; Pred. No. 31;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 ERAEQRLKSODL 13  
DB 63 ERONLQRYRSODL 75

Search completed: March 13, 2003, 18:53:53  
Job time : 12.5818 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:45:22 ; Search time 20.7273 Seconds  
(without alignments)  
99.409 Million cell updates/sec

Title: US-10-050-200-6  
Perfect score: 45  
Sequence: 1 EKARVLAEM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP. archaea: \*  
2: SP. bacteria: \*  
3: SP. fungi: \*  
4: SP. human: \*  
5: SP. invertebrate: \*  
6: SP. mammal: \*  
7: SP. mhc: \*  
8: SP. organelle: \*  
9: SP. phage: \*  
10: SP. plant: \*  
11: SP. rodent: \*  
12: SP. virus: \*  
13: SP. vertebrate: \*  
14: SP. unclassified: \*  
15: SP. rvirus: \*  
16: SP. bacteriophage: \*  
17: SP. archaea: \*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	88.9	46	15	09J122 human immun
2	40	88.9	54	15	09J181 human immun
3	40	88.9	55	15	09J140 human immun
4	40	88.9	56	15	09J117 human immun
5	40	88.9	58	15	09J144 human immun
6	40	88.9	58	15	09J138 human immun
7	40	88.9	61	15	09J198 human immun
8	40	88.9	61	15	09J197 human immun
9	40	88.9	61	15	09J168 human immun
10	40	88.9	62	15	09J132 human immun
11	40	88.9	62	15	09J612 human immun
12	40	88.9	63	15	09J177 human immun
13	40	88.9	63	15	09J126 human immun
14	40	88.9	63	15	09J123 human immun
15	40	88.9	64	15	09J189 human immun
16	40	88.9	64	15	09J184 human immun

17	40	88.9	64	15	09J182 human immun
18	40	88.9	64	15	09J167 human immun
19	40	88.9	64	15	09J159 human immun
20	40	88.9	64	15	09J155 human immun
21	40	88.9	65	15	09J195 human immun
22	40	88.9	65	15	09J174 human immun
23	40	88.9	65	15	09J171 human immun
24	40	88.9	65	15	09J160 human immun
25	40	88.9	65	15	09J158 human immun
26	40	88.9	65	15	09J147 human immun
27	40	88.9	65	15	09J139 human immun
28	40	88.9	65	15	09J137 human immun
29	40	88.9	65	15	09J136 human immun
30	40	88.9	65	15	09J120 human immun
31	40	88.9	65	15	09J115 human immun
32	40	88.9	66	15	09J186 human immun
33	40	88.9	66	15	09J175 human immun
34	40	88.9	66	15	09J173 human immun
35	40	88.9	66	15	09J127 human immun
36	40	88.9	66	15	09J125 human immun
37	40	88.9	66	15	09J597 human immun
38	40	88.9	67	15	09J194 human immun
39	40	88.9	67	15	09J176 human immun
40	40	88.9	67	15	09J163 human immun
41	40	88.9	67	15	09J152 human immun
42	40	88.9	67	15	09J142 human immun
43	40	88.9	67	15	09J141 human immun
44	40	88.9	67	15	09J129 human immun
45	40	88.9	67	15	09J128 human immun

## ALIGNMENTS

### RESULT 1

09J122 PRELIMINARY; PRT; 46 AA.  
AC 09J122;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE Gag protein (Fragment).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FIN92168;  
RA Laitola K., Holmstrom P., Laukkanen T., Brunner-Korvenkontio H.,  
RA Leinikki P., Salminen M.O.;  
RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
RT correlation between molecular and epidemiological data."  
RU Scand. J. Infect. Dis. 0:0-0(2000).  
DR EMBL: AF219351; AAF30257.1; -;  
FT NON\_TER  
SQ SEQUENCE 46 AA; 4946 MW; 7F85C2C145AF1677 CRC64;

Query Match 88.9%; Score 40; DB 15; Length 46;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARYLAEM 10  
Db 6 KARYLAEM 14

### RESULT 2

09J181 PRELIMINARY; PRT; 54 AA.  
AC 09J181;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Gag protein (Fragment).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FIN9140;  
 RA Lillsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
 RL Leinikki P., Salmunen M.O.;  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data."  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL; AF219293; AAF30198.1; -.  
 DR HSSP; P05888; IAAE.  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; zf-CCHC; 1.  
 DR SMART; SM00343; Znf\_C2HC; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 1 1  
 SQ SEQUENCE 54 AA; 6000 MW; 94C21F799195F4NC CRC64;  
 Query Match 88.9%; Score 40; DB 15; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC 09J117;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DE 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Gag protein (Fragment).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FIN9210;  
 RA Lillsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
 RL Leinikki P., Salmunen M.O.;  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data."  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL; AF219356; AAF30262.1; -.  
 DR HSSP; P05888; IAAE.  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; zf-CCHC; 1.  
 DR SMART; SM00343; Znf\_C2HC; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 1 1  
 SQ SEQUENCE 56 AA; 6191 MW; 0DA57152AD3E4482 CRC64;  
 Query Match 88.9%; Score 40; DB 15; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
 DB 5 KARVLAEM 13

## RESULT 3

ID 09J140 PRELIMINARY; PRT; 55 AA.  
 AC 09J140;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Gag protein (Fragment).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FIN907;  
 RA Lillsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
 RL Leinikki P., Salmunen M.O.;  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data."  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL; AF219333; AAF30239.1; -.  
 DR HSSP; P05888; IAAE.  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; zf-CCHC; 1.  
 DR SMART; SM00343; Znf\_C2HC; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 1 1  
 SQ SEQUENCE 55 AA; 6077 MW; 2729CB40CC66804A CRC64;  
 Query Match 88.9%; Score 40; DB 15; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
 DB 6 KARVLAEM 14

RESULT 4  
 09J117 PRELIMINARY; PRT; 56 AA.  
 ID 09J117

AC 09J144  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DE 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Gag protein (Fragment).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FIN9206;  
 RA Lillsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
 RL Leinikki P., Salmunen M.O.;  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data."  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL; AF219329; AAF30235.1; -.  
 DR HSSP; P05888; IAAE.  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; zf-CCHC; 1.  
 DR SMART; SM00343; Znf\_C2HC; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 1 1  
 SQ SEQUENCE 58 AA; 6414 MW; 86BA5037A48AE958 CRC64;  
 Query Match 88.9%; Score 40; DB 15; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
 DB 6 KARVLAEM 14

```

RESULT 6
09J138
ID 09J138 PRELIMINARY; PRT; 58 AA.
AC 09J138;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Gag protein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FIN93173;
RA Litsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,
RA Leinikki P., Salminen M.O.;
RT "Analysis of HIV-1 genetic subtypes in Finland reveals good
RT correlation between molecular and epidemiological data.";
RL Scand. J. Infect. Dis. 0:0-0(2000).
DR EMBL: AF219335; AAF30241.1; -.
DR HSSP: P05888; IAAF.
DR InterPro: IPR001878; znf_CCHC.
DR Pfam: PF00098; zf-CCHC; 1.
DR SMART: SM00343; znf_C2HC; 1.
FT NON_TER 1 1
FT NON_TER 58 58
SQ SEQUENCE 58 AA; 6313 MW; F3A2592D06632B69 CRC64;

Query Match
Best Local Similarity 100.0%; Score 40; DB 15; Length 58;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KARVLAEM 10
DB 6 KARVLAEM 14

RESULT 7
09J198
ID 09J198 PRELIMINARY; PRT; 61 AA.
AC 09J198;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Gag protein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FIN93214;
RA Litsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,
RA Leinikki P., Salminen M.O.;
RT "Analysis of HIV-1 genetic subtypes in Finland reveals good
RT correlation between molecular and epidemiological data.";
RL Scand. J. Infect. Dis. 0:0-0(2000).
DR EMBL: AF219275; AAF30181.1; -.
DR HSSP: P05888; IAAF.
DR InterPro: IPR001878; znf_CCHC.
DR Pfam: PF00098; zf-CCHC; 1.
DR SMART: SM00343; znf_C2HC; 1.
FT NON_TER 1 1
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6951 MW; D2A2F472E0ECF696 CRC64;

Query Match
Best Local Similarity 100.0%; Score 40; DB 15; Length 61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KARVLAEM 10
DB 6 KARVLAEM 14

```

```

DB 1 KARVLAEM 9

RESULT 8
09J197
ID 09J197 PRELIMINARY; PRT; 61 AA.
AC 09J197;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Gag protein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FIN93180;
RA Litsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,
RA Leinikki P., Salminen M.O.;
RT "Analysis of HIV-1 genetic subtypes in Finland reveals good
RT correlation between molecular and epidemiological data.";
RL Scand. J. Infect. Dis. 0:0-0(2000).
DR EMBL: AF219276; AAF30182.1; -.
DR HSSP: P05888; IAAF.
DR InterPro: IPR001878; znf_CCHC.
DR Pfam: PF00098; zf-CCHC; 1.
DR SMART: SM00343; znf_C2HC; 1.
FT NON_TER 1 1
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6924 MW; EC89D5870B733B92 CRC64;

Query Match
Best Local Similarity 100.0%; Score 40; DB 15; Length 61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KARVLAEM 10
DB 4 KARVLAEM 12

RESULT 9
09J168
ID 09J168 PRELIMINARY; PRT; 61 AA.
AC 09J168;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Gag protein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FIN9213;
RA Litsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,
RA Leinikki P., Salminen M.O.;
RT "Analysis of HIV-1 genetic subtypes in Finland reveals good
RT correlation between molecular and epidemiological data.";
RL Scand. J. Infect. Dis. 0:0-0(2000).
DR EMBL: AF219305; AAF30211.1; -.
DR HSSP: P05888; IAAF.
DR InterPro: IPR001878; znf_CCHC.
DR Pfam: PF00098; zf-CCHC; 1.
DR SMART: SM00343; znf_C2HC; 1.
FT NON_TER 1 1
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6932 MW; 5980BCB1D2469B1C CRC64;

Query Match
Best Local Similarity 100.0%; Score 40; DB 15; Length 61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KARVLAEM 10
DB 4 KARVLAEM 12

```

OY 2 KARVLAEM 10  
 Db 4 KARVLAEM 12

## RESULT 10

09J132 PRELIMINARY; PRT; 62 AA.  
 AC 09J132;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Gag protein (Fragment).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FIN9307;  
 RA Laitola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
 Leinikki P., Salmien M.O.;  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data."  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL: AF219341; AAF30247.1;  
 DR HSSP: P05888; IAAF.  
 DR InterPro: IPR001878; znf\_CCHC.  
 DR Pfam: PF00098; zfc\_CCHC; 1.  
 DR SMART: SM00343; znf\_C2HC; 1.  
 FT NON\_TER 1  
 FT SEQUENCE 62 AA; 6958 MW; 0E61D6E3C5197CDD CRC64;

Query Match 88.9%; Score 40; DB 15; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 Db 2 KARVLAEM 10

## RESULT 11

09J612 PRELIMINARY; PRT; 62 AA.  
 AC 09J612;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE P24/NCp7 (Fragment).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RUS9419;  
 RA Laitola K., Laukkanen T., Denisova A., Grishkevichius A.,  
 Smol'skaja T., Usina V., Vlasov N., Leinikki P., Salmien M.;  
 RT "Genetic characterization of HIV-1 strains in the Baltic Countries and  
 Russia."  
 RL Scand. J. Infect. Dis. 0:0-0(1996).  
 DR EMBL: U70301; AAB09019.1;  
 DR HSSP: P05888; IAAF.  
 DR InterPro: IPR001878; znf\_CCHC.  
 DR Pfam: PF00098; zfc\_CCHC; 1.  
 DR SMART: SM00343; znf\_C2HC; 1.  
 FT NON\_TER 1  
 FT SEQUENCE 62 AA; 6982 MW; E7D3F40AFB94B4A1 CRC64;

Query Match 88.9%; Score 40; DB 15; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 Db 1 KARVLAEM 9

## RESULT 12

09J177 PRELIMINARY; PRT; 63 AA.  
 AC 09J177;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Gag protein (Fragment).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FIN93211;  
 RA Laitola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
 Leinikki P., Salmien M.O.;  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data."  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL: AF219296; AAF30202.1;  
 DR HSSP: P05888; IAAF.  
 DR InterPro: IPR001878; znf\_CCHC.  
 DR Pfam: PF00098; zfc\_CCHC; 1.  
 DR SMART: SM00343; znf\_C2HC; 1.  
 FT NON\_TER 1  
 FT SEQUENCE 63 AA; 7133 MW; 3D6A185BC59A13D8 CRC64;

Query Match 88.9%; Score 40; DB 15; Length 63;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 Db 2 KARVLAEM 10

## RESULT 13

09J126 PRELIMINARY; PRT; 63 AA.  
 AC 09J126;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Gag protein (Fragment).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FIN9401;  
 RA Laitola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
 Leinikki P., Salmien M.O.;  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data."  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL: AF219347; AAF30253.1;  
 DR HSSP: P05888; IAAF.  
 DR InterPro: IPR001878; znf\_CCHC.  
 DR Pfam: PF00098; zfc\_CCHC; 1.  
 DR SMART: SM00343; znf\_C2HC; 1.  
 FT NON\_TER 1

FT NON\_TER 63 63  
SQ SEQUENCE 63 AA; 6922 MW; 4CB95C8F39BD5394 CRC64;

Query Match  
Best Local Similarity 88.9%; Score 40; DB 15; Length 63;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
DB 6 KARVLAEM 14

## RESULT 14

O9J123 PRELIMINARY; PRT; 63 AA.

AC O9J123;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Gag protein (Fragment).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovirdae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FIN9149;  
RA Liitsola K., Holmstrom P., Laukkanen T., Brunner-Korvenkontio H.,  
RA Leinikki P., Salminen M.O.;  
RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
RT correlation between molecular and epidemiological data.";  
RL Scand. J. Infect. Dis 0:0-0(2000).  
DR EMBL: AF219350; AAF30256.1; -;  
DR HSSP: P05888; 1AAF.  
DR InterPro: IPR001878; Znf\_CCHC.  
DR Pfam: PF00098; zf-CCHC; 1.  
DR SMART: SM00343; Znf\_C2HC; 1.  
FT NON\_TER 1  
FT NON\_TER 63  
SQ SEQUENCE 63 AA; 7078 MW; D82DF25E049C42D0 CRC64;

Query Match 88.9%; Score 40; DB 15; Length 63;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
DB 4 KARVLAEM 12

## RESULT 15

O9J189 PRELIMINARY; PRT; 64 AA.

AC O9J189;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Gag protein (Fragment).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrovirdae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FIN93230;  
RA Liitsola K., Holmstrom P., Laukkanen T., Brunner-Korvenkontio H.,  
RA Leinikki P., Salminen M.O.;  
RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
RT correlation between molecular and epidemiological data.";  
RL Scand. J. Infect. Dis 0:0-0(2000).  
DR EMBL: AF219284; AAF30190.1; -;  
DR HSSP: P05888; 1AAF.  
DR InterPro: IPR001878; Znf\_CCHC.

DR Pfam: PF00098; zf-CCHC; 1.  
DR SMART: SM00343; Znf\_C2HC; 1.

FT NON\_TER 1  
FT NON\_TER 64  
SQ SEQUENCE 64 AA; 7117 MW; FD369389B20186AC CRC64;

Query Match 88.9%; Score 40; DB 15; Length 64;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
DB 2 KARVLAEM 10

Search completed: March 13, 2003, 18:51:57  
Job time : 20.7273 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:40:28 ; Search time 5.27273 Seconds  
(without alignments)  
78.662 Million cell updates/sec

Title: US-10-050-200-6  
Perfect score: 45  
Sequence: 1 EKARYLAEM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	88.9	388	1 GAG_HV1W2	P05889 human immun
2	40	88.9	492	1 GAG_HV1U4	P24736 human immun
3	40	88.9	496	1 GAG_HV1ND	P18800 human immun
4	40	88.9	498	1 GAG_HV1OY	P20889 human immun
5	40	88.9	499	1 GAG_HV1C4	P05887 human immun
6	40	88.9	499	1 GAG_HV1E1	P04592 human immun
7	40	88.9	499	1 GAG_HV1H2	P04591 human immun
8	40	88.9	499	1 GAG_HV1J3	P12494 human immun
9	40	88.9	499	1 GAG_HV1M	P07022 human immun
10	40	88.9	499	1 GAG_HV1N5	P12493 human immun
11	40	88.9	499	1 GAG_HV1Y2	P35862 human immun
12	40	88.9	500	1 GAG_HV1Z2	P12495 human immun
13	40	88.9	501	1 GAG_HV1A2	P03349 human immun
14	40	88.9	503	1 GAG_HV1JR	P20873 human immun
15	40	88.9	504	1 GAG_HV1MA	P04594 human immun
16	40	88.9	506	1 GAG_HV1MN	P05888 human immun
17	40	88.9	508	1 GAG_HV1VZ	P17282 chimpanzee
18	40	88.9	511	1 GAG_HV1B1	P03347 human immun
19	40	88.9	511	1 GAG_HV1B5	P04593 human immun
20	40	88.9	511	1 GAG_HV1BR	P03348 human immun
21	40	88.9	511	1 GAG_HV1PV	P03350 human immun
22	39	86.7	500	1 GAG_HV1RH	P05890 human immun
23	38	84.4	298	1 ARGB_NE1MA	Q9JUK2 neisseria m
24	38	84.4	298	1 ARGB_NE1MB	Q9JUK2 neisseria m
25	36	80.0	283	1 AROE_METTH	O26344 methanobact
26	36	80.0	306	1 BLAC_STRFR	P35392 streptomyc
27	34	75.6	298	1 ARGB_AQUAE	O67848 aquifex aeo
28	34	75.6	340	1 TRPD_BACPU	P18261 bacillus pu
29	34	75.6	506	1 GAG_STVM1	P05894 bacillus immu
30	34	75.6	506	1 GAG_STVMK	P05893 simian immu
31	34	75.6	507	1 GAG_STVSK	P12496 simian immu
32	34	75.6	507	1 GAG_STVSP	P19504 simian immu
33	34	75.6	510	1 GAG_STVMS	P31634 simian immu

34	34	75.6	519	1 GAG_HV2NZ	P05891 human immun
35	34	75.6	520	1 GAG_HV2SB	P12450 human immun
36	34	75.6	521	1 GAG_HV2BE	P18095 human immun
37	34	75.6	521	1 GAG_HV2CA	P24106 human immun
38	34	75.6	521	1 GAG_HV2D1	P17756 human immun
39	34	75.6	521	1 GAG_HV2D2	P15832 human immun
40	34	75.6	521	1 GAG_HV2KR	O74119 human immun
41	34	75.6	521	1 GAG_HV2ST	P20874 human immun
42	34	75.6	522	1 GAG_HV2G1	P18041 human immun
43	34	75.6	522	1 GAG_HV2RO	P04590 human immun
44	34	75.6	557	1 YP85_MYCTU	Q0636 mycobacteri
45	34	75.6	785	1 MUS2_BACHD	O9K8A0 bacillus ha

## ALIGNMENTS

RESULT 1  
ID GAG\_HV1W2 STANDARD; PRT; 388 AA.  
AC P05889;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE GAG polyprotein [contains: Core proteins p17, p24, p2, p7, p1, p6]  
DE (Fragment).  
GN GAG.  
OS Human Immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11705;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86235450; PubMed=3012778;  
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,  
RA Salathuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;  
RT "Genetic variation in HIV-1/LAV over time in patients with AIDS or  
RT at risk for AIDS";  
RT Science 232:1548-1553(1986).  
RL  
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM  
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO  
CC WAS PERMANENTLY INFECTED BY HER MOTHER.  
CC  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: K03457; AAB12988.1; -  
CC DR HSSP: P03351; IE1A.  
DR HIV: K03457; GAG\$WMJ2.  
DR InterPro: IPR000721; Gag\_p24.  
DR InterPro: IPR000071; Retroviral\_p17.  
DR Pfam: PF00540; Gag\_p17; 1.  
DR Pfam: PF00607; Gag\_p24; 1.  
KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;  
KW Zinc-finger.  
FT INIT\_MET 0  
FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).  
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).  
FT CHAIN 363 376 CORE PROTEIN P2.  
FT CHAIN 377 >388 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).

```

FT  NON_TER      388      388
SQ  SEQUENCE      388 AA: 43314 MM: EFB85BE10ECF7804 CRC64:
      Query Match
      Best Local Similarity 100.0%; Score 40; DB 1; Length 388;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2 KARVLAEM 10
      |||||
      358 KARVLAEM 366

RESULT 2
GAG_HV104      STANDARD;      PRT;      492 AA.
ID  GAG_HV104
AC  P24736;
DT  01-MAR-1992 (Rel. 21, Created)
DT  01-FEB-1994 (Rel. 28, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  GAG polypeptide [Contains: Core proteins p17, p24, p7, p1, p6].
GN  GAG.
OS  Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)
OC  Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX  NCBI_TaxID=11703;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=91090981; PubMed=2265025;
RA  Oram J.D., Downing R.G., Rolf M., Clegg J.C.S., Serwadda D.,
RA  Carswell J.W.;
RT  "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic
RT  diversity from other HIV-1 isolates."
RL  AIDS Res. Hum. Retroviruses 6:1073-1078(1990).
CC  -1- FUNCTION: PERFORMS HIGHLY ORCHESTRATED TASKS DURING THE
CC  ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC  REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC  MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC  RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC  GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC  BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC  -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.
CC  -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL: M62320; AAA75018.1; -
DR  HSSP: P05888; IAA7.
DR  InterPro: IPR000721; Gag_p24.
DR  InterPro: IPR000071; Retrovir_p17.
DR  InterPro: IPR001878; Znf_CCHC.
DR  Pfam: PF00098; zf_CCHC; 2.
DR  Pfam: PF00540; Gag_p17; 1.
DR  Pfam: PF00607; Gag_p24; 1.
DR  PRINTS: PR00939; C2HCZNFINGER.
DR  PRINTS: PR00234; HIV1MATRIX.
DR  SMART: SM00343; Znf_C2HC; 2.
DR  PROSITE: PS50158; ZF_CCHC; 2.
KW  AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
KW  Zinc-finger; Repeat.
FT  INT_MET      0
FT  CHAIN      1 127      BY SIMILARITY.
FT  CHAIN      128 358      CORE PROTEIN P17 (MATRIX PROTEIN).
FT  CHAIN      359 371      CORE PROTEIN P24 (CORE ANTIGEN).
FT  CHAIN      372 425      CORE PROTEIN P2.
FT  CHAIN      426 441      CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT  CHAIN      442 492      CORE PROTEIN P1.
FT  CHAIN      493 492      CORE PROTEIN P6.
FT  ZN_FING      383 400      CCHC-TYPE 1.

```

```

FT  ZN_FING      404 421      CCHC-TYPE 2.
FT  LIPID      1
SQ  SEQUENCE      492 AA: 54791 MM: MYRISTATE (BY SIMILARITY).
      Query Match
      Best Local Similarity 100.0%; Score 40; DB 1; Length 492;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2 KARVLAEM 10
      |||||
      354 KARVLAEM 362

RESULT 3
GAG_HV1ND      STANDARD;      PRT;      496 AA.
ID  GAG_HV1ND
AC  P18800;
DT  01-NOV-1990 (Rel. 16, Created)
DT  01-FEB-1994 (Rel. 28, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  GAG polypeptide [Contains: Core proteins p17, p24, p7, p1, p6].
GN  GAG.
OS  Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC  Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX  NCBI_TaxID=11695;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=90034200; PubMed=2806917;
RA  Spire B., Site J., Zachar V., Rey F., Barre-Sinoussi F., Gallibert F.,
RA  Hampe A., Chermann J.C.;
RT  "Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the
RT  human immunodeficiency virus."
RL  Gene 81:275-284(1989).
CC  -1- FUNCTION: PERFORMS HIGHLY ORCHESTRATED TASKS DURING THE
CC  ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC  REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC  MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC  RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC  GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC  BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC  -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.
CC  -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC  AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC  -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL: M27323; AAA44868.1; -
DR  PIR: J00065; FOLIND.
DR  HSSP: P05888; IAA7.
DR  HIV: M27323; GAGSNDK.
DR  InterPro: IPR000721; Gag_p24.
DR  InterPro: IPR000071; Retrovir_p17.
DR  InterPro: IPR001878; Znf_CCHC.
DR  Pfam: PF00098; zf_CCHC; 2.
DR  Pfam: PF00540; Gag_p17; 1.
DR  Pfam: PF00607; Gag_p24; 1.
DR  PRINTS: PR00939; C2HCZNFINGER.
DR  PRINTS: PR00234; HIV1MATRIX.
DR  SMART: SM00343; Znf_C2HC; 2.
DR  PROSITE: PS50158; ZF_CCHC; 2.
KW  AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
KW  Zinc-finger; Repeat.
FT  INT_MET      0
FT  CHAIN      1 128      BY SIMILARITY.
FT  CHAIN      129 359      CORE PROTEIN P17 (MATRIX PROTEIN).
FT  CHAIN      360 374      CORE PROTEIN P24 (CORE ANTIGEN).

```



FT CHAIN 375 429 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
 FT CHAIN 430 445 CORE PROTEIN P1.  
 FT CHAIN 446 496 CORE PROTEIN P6.  
 FT 2N\_FING 387 404 CCHC-TYPE 1.  
 FT 2N\_FING 408 425 CCHC-TYPE 2.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 SQ SEQUENCE 496 AA; 55152 MW; F50A42E42CCF20AA CRC64;  
 Query Match 88.9%; Score 40; DB 1; Length 496;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KARVLAEM 10  
 Db 355 KARVLAEM 363  
 RESULT 4  
 GAG\_HV10Y STANDARD; PRT; 498 AA.  
 AC P20889;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (HTI isolate) (HIV-1).  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11699;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90148544; PubMed=2559749;  
 RA Huet T., Darza M.C., Brun-Vezinet F., Roelants G.E., Main-Hobson S.;  
 RT "A highly defective HIV-1 strain isolated from a healthy Gabonese  
 individual presenting an atypical western blot."  
 RL AIDS 3:707-715(1989).  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- MISCELLANEOUS: THE HTI ISOLATE WAS TAKEN FROM THE BLOOD OF A  
 HEALTHY GABONESE INDIVIDUAL.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M26727; AAA8391.1; -  
 CC HSSP; P05888; IAAF.  
 DR HIV; M26727; GAGSOYI.  
 DR InterPro: IPR000721; Gag\_P24.  
 DR InterPro: IPR000071; Retrovir\_p17.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; Zf-CCHC; 2.  
 DR Pfam; PF00540; Gag\_P17; 1.  
 DR Pfam; PF00639; Gag\_P24; 1.  
 DR PRINTS; PR00939; C2HCZNFINGER.  
 DR PRINTS; PR00234; HIVMATRIX.  
 DR SMART; SM00343; ZNF\_CCHC; 2.  
 DR PROSITE; PS50158; ZF\_CCHC; 2.  
 DR AIDS; Core protein; Polypotein; Myristate; Phosphorylation;  
 KW ZINC-finger; Repeat.  
 FT INIT MET 0  
 FT CHAIN 1 131 BY SIMILARITY.  
 CORE PROTEIN P17 (MATRIX PROTEIN).

FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).  
 FT CHAIN 363 376 CORE PROTEIN P2.  
 FT CHAIN 377 430 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
 FT CHAIN 431 446 CORE PROTEIN P1.  
 FT CHAIN 447 498 CORE PROTEIN P6.  
 FT 2N\_FING 388 405 CCHC-TYPE 1.  
 FT 2N\_FING 409 426 CCHC-TYPE 2.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 SQ SEQUENCE 498 AA; 55451 MW; BDF20658DEB20B1 CRC64;  
 Query Match 88.9%; Score 40; DB 1; Length 498;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KARVLAEM 10  
 Db 358 KARVLAEM 366  
 RESULT 5  
 GAG\_HV1C4 STANDARD; PRT; 499 AA.  
 AC P05887;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (CDC-451 isolate) (HIV-1).  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11687;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87041461; PubMed=3490666;  
 RA Desai S.M., Kalyanaraman V.S., Casey J.M., Srinivasan A.,  
 RA Andersen P.R., Devare S.G.;  
 RT "Molecular cloning and primary nucleotide sequence analysis of a  
 distinct human immunodeficiency virus isolate reveal significant  
 divergence in its genomic sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:8380-8384(1986).  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M13136; AAA44306.1; -  
 CC PIR; A25523; FOYVHA.  
 DR HSSP; P05888; IAAF.  
 DR HIV; M13136; GAGSCDC45.  
 DR InterPro: IPR000721; Gag\_P24.  
 DR InterPro: IPR000071; Retrovir\_p17.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; Zf-CCHC; 2.  
 DR Pfam; PF00540; Gag\_P17; 1.  
 DR Pfam; PF00607; Gag\_P24; 1.  
 DR PRINTS; PR00939; C2HCZNFINGER.  
 DR PRINTS; PR00234; HIVMATRIX.  
 DR SMART; SM00343; ZNF\_CCHC; 2.  
 DR PROSITE; PS50158; ZF\_CCHC; 2.  
 DR AIDS; Core protein; Polypotein; Myristate; Phosphorylation;

KW Zinc-finger: Repeat. 0  
 FT INIT\_MER 0  
 FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).  
 FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).  
 FT CHAIN 363 376 CORE PROTEIN P2.  
 FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
 FT CHAIN 432 447 CORE PROTEIN P1.  
 FT CHAIN 448 499 CORE PROTEIN P6.  
 FT ZN\_FING 389 406 CCHC-TYPE 1.  
 FT ZN\_FING 410 427 CCHC-TYPE 2.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 SQ SEQUENCE 499 AA: 55796 MW: 023CA76C9C6F22AD CRC64;

Query Match 88.9%; Score 40; DB 1; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 DB 358 KARVLAEM 366

RESULT 6  
 GAG\_HV1EL STANDARD; PRT; 499 AA.  
 ID GAG\_HV1EL  
 AC P04592;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid-11689;  
 RN [1]  
 RP MEDLINE-86245056; PubMed-2424612;  
 RA Alton M., Main-Hobson S., Montagnier L., Sonigo P.;  
 RT "Genetic variability of the AIDS virus: nucleotide sequence analysis  
 of two isolates from African patients.";  
 RL Cell 46:63-74(1986).  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
 or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
 CC -----  
 DR EMBL: K03454; AAA4324.1; -;  
 DR EMBL: A07108; CA000611.1; -;  
 DR HSSP: P05888; IAAF.  
 DR HIV: K03454; GAGSEL1.  
 DR InterPro: IPR000721; Gag\_P24.  
 DR InterPro: IPR000071; Retrovir\_P17.  
 DR InterPro: IPR001878; znf\_CCHC.  
 DR Pfam: PF00098; zf\_CCHC; 2.  
 DR Pfam: PF00540; Gag\_P17; 1.  
 DR Pfam: PF00607; Gag\_P24; 1.  
 DR PRINTS: PR00939; C2HCZNFINGER.  
 DR PRINTS: PR00234; HIVMATRIX.  
 DR SMART: SM00343; znf\_C2HC; 2.  
 DR PROSITE: PS50158; znf\_CCHC; 2.

KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;  
 KW Zinc-finger: Repeat. 0  
 FT INIT\_MER 0  
 FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).  
 FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).  
 FT CHAIN 363 377 CORE PROTEIN P2.  
 FT CHAIN 378 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
 FT CHAIN 432 447 CORE PROTEIN P1.  
 FT CHAIN 448 499 CORE PROTEIN P6.  
 FT ZN\_FING 390 407 CCHC-TYPE 1.  
 FT ZN\_FING 411 428 CCHC-TYPE 2.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 SQ SEQUENCE 499 AA: 55562 MW: 8A1785A59EAE08D CRC64;

Query Match 88.9%; Score 40; DB 1; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 DB 358 KARVLAEM 366

RESULT 7  
 GAG\_HV1H2 STANDARD; PRT; 499 AA.  
 ID GAG\_HV1H2  
 AC P04591;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid-11706;  
 RN [1]  
 RP MEDLINE-87299196; PubMed-3040055;  
 RA Ratner L., Fisher A., Jagodzinski L., Mitsuya H., Liou R.-S.,  
 RA Gallo R.C., Wong-Staal F.;  
 RT "Complete nucleotide sequences of functional clones of the AIDS  
 virus.";  
 RL AIDS Res. Hum. Retroviruses 3:57-69(1987).  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
 or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
 CC -----  
 DR EMBL: K03455; AAB50258.1; -;  
 DR HSSP: P05888; IAAF.  
 DR HIV: K03455; GAGSHX2.  
 DR InterPro: IPR000721; Gag\_P24.  
 DR InterPro: IPR001878; znf\_CCHC.  
 DR Pfam: PF00098; zf\_CCHC; 2.  
 DR Pfam: PF00540; Gag\_P17; 1.  
 DR Pfam: PF00607; Gag\_P24; 1.  
 DR PRINTS: PR00939; C2HCZNFINGER.  
 DR PRINTS: PR00234; HIVMATRIX.  
 DR SMART: SM00343; znf\_C2HC; 2.

```

DR PROSITE: PS50158; 2F_CCHC; 2.
KM AIDS; Core protein: Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat.
FT INIT_MET 0 0 BY SIMILARITY.
FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 363 376 CORE PROTEIN P2.
FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 432 447 CORE PROTEIN P6.
FT CHAIN 448 499 CORE PROTEIN P1.
FT ZN_FING 389 406 CCHC-TYPE 1.
FT ZN_FING 410 427 CCHC-TYPE 2.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 499 AA; 55798 MW; 774C384D6EACB108 CRC64;

Query Match 88.9%; Score 40; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10
DB 358 KARVLAEM 366

RESULT 8
GAG_HVLJ3 STANDARD; PRT; 499 AA.
AC P12494;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (H3 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11694;
RN [1]
RP MEDLINE=89352108; PubMed=2669897;
RA Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;
RT "Nucleotide sequences of gag and env genes of a Japanese isolate of HIV-1 and their expression in bacteria.";
RL AIDS Res. Hum. Retroviruses 5:411-419(1989).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL. GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: M2137; AAB03522.1; -
CC HSSP: P05888; IAAF.
DR HIV; M2137; GAG$JH3.
DR InterPro: IPR000721; Gag_p24.
DR InterPro: IPR000071; Retrovir_p17.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00098; zf-CCHC; 2.
DR Pfam: PF00540; Gag_p17; 1.
DR Pfam: PF00607; Gag_p24; 1.
DR PRINTS: PR00939; C2HCZNFINGER.
DR PRINTS: PR00234; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.

```

```

DR PROSITE: PS50158; 2F_CCHC; 2.
KM AIDS; Core protein: Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat.
FT INIT_MET 0 0 BY SIMILARITY.
FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 363 376 CORE PROTEIN P2.
FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 432 447 CORE PROTEIN P6.
FT CHAIN 448 499 CORE PROTEIN P1.
FT ZN_FING 389 406 CCHC-TYPE 1.
FT ZN_FING 410 427 CCHC-TYPE 2.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 499 AA; 55725 MW; 52E281D4424AE9 CRC64;

Query Match 88.9%; Score 40; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10
DB 358 KARVLAEM 366

RESULT 9
GAG_HVILM STANDARD; PRT; 499 AA.
AC Q70622;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (HM12.3 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=82834;
RN [1]
RP MEDLINE=95127297; PubMed=7826699;
RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M., Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type IIIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL. GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: U12055; AAA76686.1; -
CC HSSP: P05888; IAAF.
DR HIV; U12055; GAG$JH3.
DR InterPro: IPR000721; Gag_p24.
DR InterPro: IPR000071; Retrovir_p17.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00098; zf-CCHC; 2.
DR Pfam: PF00540; Gag_p17; 1.
DR Pfam: PF00607; Gag_p24; 1.
DR PRINTS: PR00939; C2HCZNFINGER.
DR PRINTS: PR00234; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.

```

DR PROSITE: PS50158; ZF\_CCHC; 2.  
 KM AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;  
 KW Zinc-finger; Repeat.

FT	INIT_MET	0	BY SIMILARITY.
FT CHAIN	1	131	CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN	132	362	CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN	363	376	CORE PROTEIN P2.
FT CHAIN	377	431	CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN	432	447	CORE PROTEIN P1.
FT CHAIN	448	499	CORE PROTEIN P6.
FT ZN_FING	389	406	CCHC-TYPE 1.
FT ZN_FING	410	427	CCHC-TYPE 2.
FT LIPID	1	1	MYRISTATE (BY SIMILARITY).
SEQUENCE	499 AA:	55772 MW:	AA2F683546EDC0A9 CRC64;

Query Match 88.9%; Score 40; DB 1; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 |||||  
 Db 358 KARVLAEM 366

RESULT 10  
 GAG\_HV1N5 STANDARD; PRT; 499 AA.  
 AC P12493;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polyprotein [Contains: Core proteins p17, p24, p2, p7, p1, p6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11698;  
 RN [1]  
 RP SEQUENCE FROM N.A. (CLONE PNL4-3).  
 RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;  
 RL Submitted (JUN-1988) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: PERFORMS HIGHLY ORCHESTRATED TASKS DURING THE  
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M19921; AAA44987.1; -  
 DR HSSP: P05888; 1A4F.  
 DR HIV: M19921; GAGS1A3.  
 DR InterPro: IPR000721; Gag\_p24.  
 DR InterPro: IPR000721; Retrovir\_p17.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam: PF00098; ZF\_CCHC; 2.  
 DR Pfam: PF00540; Gag\_P17; 1.  
 DR Pfam: PF00607; Gag\_P24; 1.  
 DR PRINTS: PR00939; C2HCZNFINGER.  
 DR PRINTS: PR00234; HIV1MATRIX.  
 DR SMART: SM00343; ZNF\_C2HC; 2.  
 DR PROSITE: PS50158; ZF\_CCHC; 2.  
 KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;  
 ZNF-finger; Repeat.

FT INIT\_MET 0 0 BY SIMILARITY.  
 FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).  
 FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).  
 FT CHAIN 363 376 CORE PROTEIN P2.  
 FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
 FT CHAIN 432 447 CORE PROTEIN P1.  
 FT CHAIN 448 499 CORE PROTEIN P6.  
 FT ZN\_FING 389 406 CCHC-TYPE 1.  
 FT ZN\_FING 410 427 CCHC-TYPE 2.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 SEQUENCE 499 AA: 55687 MW: C8EC1302FE221E2 CRC64;

Query Match 88.9%; Score 40; DB 1; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 |||||  
 Db 358 KARVLAEM 366

RESULT 11  
 GAG\_HV1Y2 STANDARD; PRT; 499 AA.  
 AC P35962;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polyprotein [Contains: Core proteins p17, p24, p2, p7, p1, p6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=36377;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-93021387; PubMed-1404605;  
 RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,  
 RA Shaw G.M.;  
 RT "Complete nucleotide sequence, genome organization, and biological  
 properties of human immunodeficiency virus type 1 in vivo: evidence  
 for limited defectiveness and complementation."  
 RL J. Virol. 66:6587-6600(1992).  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M93258; -; NOT ANNOTATED\_CDS.  
 DR PIR: A44001; A44001.  
 DR HSSP: P05888; 1A4F.  
 DR InterPro: IPR000721; Gag\_p24.  
 DR InterPro: IPR000721; Retrovir\_p17.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam: PF00098; ZF\_CCHC; 2.  
 DR Pfam: PF00540; Gag\_P17; 1.  
 DR Pfam: PF00607; Gag\_P24; 1.  
 DR PRINTS: PR00939; C2HCZNFINGER.  
 DR PRINTS: PR00234; HIV1MATRIX.  
 DR SMART: SM00343; ZNF\_C2HC; 2.  
 DR PROSITE: PS50158; ZF\_CCHC; 2.

KM AIDS; Core protein: Polypeptide; Myristate; Phosphorylation;  
 KM Zinc-finger; Repeat. 0  
 FT CHAIN 1 131 BY SIMILARITY.  
 FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).  
 FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).  
 FT CHAIN 377 441 CORE PROTEIN P2.  
 FT CHAIN 432 437 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
 FT CHAIN 448 499 CORE PROTEIN P6.  
 FT ZN\_FING 389 406 CCHC-TYPE 1.  
 FT ZN\_FING 410 427 CCHC-TYPE 2.  
 SQ SEQUENCE 499 AA; 55660 MW; 278665F5405CD99 CRC64;  
 MYRISTATE (BY SIMILARITY).  
 Query Match 88.9%; Score 40; DB 1; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 KARVLAEM 10  
 DB 358 KARVLAEM 366  
 RESULT 12  
 ID GAG\_HV1A2 STANDARD; PRT; 500 AA.  
 AC P12495;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polypeptide [contains: Core proteins P17, P24, P2, P7, P1, P6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (22/CDC-234 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11683;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Theodore T., Buckler-White A.;  
 RL Submitted (NOV-1988) to the HIV data bank.  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M22639; AAA5365.1; -.  
 DR HSSP: P05888; IAAF.  
 DR HIV: M22639; GAG52226.  
 DR InterPro: IPR000721; Gag\_p24.  
 DR InterPro: IPR000071; Retrovir\_p17.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam: PF00098; Zf-CCHC; 2.  
 DR Pfam: PF00540; Gag\_p17; 1.  
 DR Pfam: PF00607; Gag\_p24; 1.  
 DR PRINTS: PR00939; C2HCZNFINGER.  
 DR PRINTS: PR00234; HIVMATRIX.  
 DR SMART: SM00343; Znf\_CCHC; 2.  
 DR PROSITE: PS0158; Zf\_CCHC; 2.  
 KM AIDS; Core protein: Polypeptide; Myristate; Phosphorylation;  
 KM Zinc-finger; Repeat. 0  
 FT INIT\_MET 0 0 BY SIMILARITY.

FT CHAIN 1 132 CORE PROTEIN P17 (MATRIX PROTEIN).  
 FT CHAIN 133 363 CORE PROTEIN P24 (CORE ANTIGEN).  
 FT CHAIN 364 378 CORE PROTEIN P7.  
 FT CHAIN 379 433 CORE PROTEIN P2.  
 FT CHAIN 434 449 CORE PROTEIN P1.  
 FT CHAIN 450 500 CORE PROTEIN P6.  
 FT ZN\_FING 391 408 CCHC-TYPE 1.  
 FT ZN\_FING 412 429 CCHC-TYPE 2.  
 SQ SEQUENCE 500 AA; 55546 MW; B27C2609A932F82A CRC64;  
 MYRISTATE (BY SIMILARITY).  
 Query Match 88.9%; Score 40; DB 1; Length 500;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 KARVLAEM 10  
 DB 359 KARVLAEM 367  
 RESULT 13  
 ID GAG\_HV1A2 STANDARD; PRT; 501 AA.  
 AC P03349;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polypeptide [contains: Core proteins P17, P24, P2, P7, P1, P6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=85090453; PubMed=2578227;  
 RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,  
 RA Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,  
 RA Levy J.A., Dina D., Luciw P.A.;  
 RT "Nucleotide sequence and expression of an AIDS-associated retrovirus  
 (ARV-2)."  
 RL Science 227:484-492(1985).  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: K02007; AAB59875.1; -.  
 DR PIR: A03947; FOYWA2.  
 DR HSSP: P05888; IAAF.  
 DR HIV: K02007; GAG55F2.  
 DR InterPro: IPR000721; Gag\_p24.  
 DR InterPro: IPR000071; Retrovir\_p17.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam: PF00098; Zf-CCHC; 2.  
 DR Pfam: PF00540; Gag\_p17; 1.  
 DR Pfam: PF00607; Gag\_p24; 1.  
 DR PRINTS: PR00939; C2HCZNFINGER.  
 DR PRINTS: PR00234; HIVMATRIX.  
 DR SMART: SM00343; Znf\_CCHC; 2.  
 DR PROSITE: PS0158; Zf\_CCHC; 2.

```

KM AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat.
FT INIT_MET 0 0 BY SIMILARITY.
FT CHAIN 1 133 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 134 364 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 365 378 CORE PROTEIN P2.
FT CHAIN 379 433 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 434 449 CORE PROTEIN P1.
FT CHAIN 450 501 CORE PROTEIN P6.
FT ZN_FING 391 408 CCHC-TYPE 1.
FT ZN_FING 412 429 CCHC-TYPE 2.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 501 AA: 55935 MW: 8F36C928D5EDBA45 CRC64;

Query Match 88.9%; Score 40; DB 1; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10
DB 360 KARVLAEM 368

RESULT 14
GAG_HV1JR STANDARD: PRT: 503 AA.
AC P20873:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GAG polyprotein [contains: Core proteins P17, P24, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11688;
RN [1]
RP SEQUENCE FROM N.A.
RA Koyanagi S., Chen I.S.Y.;
RL Submitted (DEC-1988) to the HIV data bank.
CC -1- FUNCTION: PERFORMS HIGHLY ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M38429; AAB03744.1; -.
CC HSSP: P05888; IAAE.
CC HIV: M38429; GAGSURCSF.
CC InterPro: IPR000721; Gag_p24.
CC InterPro: IPR001878; Retrovir_P17.
CC Pfam: PF00098; zf-CCHC; 2.
CC Pfam: PF00540; Gag_P17; 1.
CC Pfam: PF00607; Gag_P24; 1.
CC PRINTS: PR00939; C2HCZNFINGER.
CC PRINTS: PR00234; HIVMATRIX.
CC SMART: SM00343; znf_C2HC; 2.
CC PROSITE: PS50158; zf_C2HC; 2.
KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat.
FT INIT_MET 0 BY SIMILARITY.

```

```

FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 363 376 CORE PROTEIN P2.
FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 432 447 CORE PROTEIN P1.
FT CHAIN 448 503 CORE PROTEIN P6.
FT ZN_FING 389 406 CCHC-TYPE 1.
FT ZN_FING 410 427 CCHC-TYPE 2.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 503 AA: 56352 MW: F0B0471CC09B42ED CRC64;

Query Match 88.9%; Score 40; DB 1; Length 503;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10
DB 358 KARVLAEM 366

RESULT 15
GAG_HV1MA STANDARD: PRT: 504 AA.
AC P04594:
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GAG polyprotein [contains: Core proteins P17, P24, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (MAL isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11697;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-86245056; PubMed-2424612;
RA Allzon M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
RT of two isolates from African patients.";
RL Cell 46:63-74(1986).
CC -1- FUNCTION: PERFORMS HIGHLY ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X04415; CA828011.1; -.
CC HSSP: P05888; IAAE.
CC HIV: K03456; GAGSMAL.
CC InterPro: IPR000721; Gag_p24.
CC InterPro: IPR000071; Retrovir_P17.
CC InterPro: IPR001878; znf_CCHC.
CC Pfam: PF00098; zf-CCHC; 2.
CC Pfam: PF00540; Gag_P17; 1.
CC Pfam: PF00607; Gag_P24; 1.
CC PRINTS: PR00939; C2HCZNFINGER.
CC PRINTS: PR00234; HIVMATRIX.
CC SMART: SM00343; znf_C2HC; 2.
CC PROSITE: PS50158; zf_C2HC; 2.
KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat.

```

	INIT_MET	0	0	BY SIMILARITY.
FT CHAIN	1	137		CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN	138	368		CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN	369	383		CORE PROTEIN P2.
FT CHAIN	384	437		CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN	438	453		CORE PROTEIN P1.
FT CHAIN	454	504		CORE PROTEIN P6.
FT ZN_FING	395	412		CCHC-TYPE 1.
FT ZN_FING	416	433		CCHC-TYPE 2.
FT LIPID	1	1		MYRISTATE (BY SIMILARITY).
SEQ SEQUENCE	504 AA;	56001 MM;	7F96785683D3BBB3 CRC64;	

Query Match 88.9%; Score 40; DB 1; length 504;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEAM 10  
DB 364 KARVLAEAM 372

Search completed: March 13, 2003, 18:49:54  
Job time : 6.27273 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:17 ; Search time 9.81818 Seconds  
(without alignments)  
97.915 Million cell updates/sec

Title: US-10-050-200-6

Sequence: 1 EKARVLAEM 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	88.9	75	2	S24478	gag polyprotein -
2	40	88.9	75	2	S24475	gag polyprotein -
3	40	88.9	75	2	S24474	gag polyprotein -
4	40	88.9	76	2	S24476	gag polyprotein -
5	40	88.9	478	1	FOVWVL	gag polyprotein -
6	40	88.9	457	1	FOLJND	gag polyprotein -
7	40	88.9	498	2	T09436	gag polyprotein -
8	40	88.9	500	1	FOVWLV	gag polyprotein -
9	40	88.9	500	1	A44001	gag polyprotein -
10	40	88.9	500	1	FOVWH4	gag polyprotein -
11	40	88.9	500	2	S33979	gag polyprotein -
12	40	88.9	501	2	S54377	gag polyprotein -
13	40	88.9	502	1	FOVWAZ	gag polyprotein -
14	40	88.9	505	2	T01667	gag polyprotein -
15	40	88.9	506	1	A38068	gag polyprotein -
16	40	88.9	508	1	FOLJST	gag polyprotein -
17	40	88.9	512	1	FOVWH3	gag polyprotein -
18	38	84.4	258	2	G81895	Probable acetylglu
19	38	84.4	258	2	E81124	acetylglutamate ki
20	37	82.2	74	2	S24473	gag polyprotein -
21	37	82.2	77	2	S24471	gag polyprotein -
22	36	80.0	260	2	C69130	gag polyprotein -
23	36	80.0	306	1	B45822	shikimate 5-dehydr
24	35	77.8	72	2	S24477	beta-lactamase (Ec
25	35	77.8	73	2	S24472	gag polyprotein -
26	35	77.8	554	2	S46346	gag polyprotein -
27	34	75.6	122	2	S72820	hypothetical prote
28	34	75.6	263	2	S72978	hypothetical prote
29	34	75.6	281	2	G86956	conserved hypotnet

#### ALIGNMENTS

30	34	75.6	340	1	JH0099	anthranilate phosph
31	34	75.6	399	2	E86289	t16N1.7 protein -
32	34	75.6	499	2	D87363	conserved hypotnet
33	34	75.6	506	1	FOLJG3	gag polyprotein -
34	34	75.6	506	1	FOLJG5	gag polyprotein -
35	34	75.6	507	2	T11559	gag protein - siml
36	34	75.6	507	2	S04237	gag polyprotein -
37	34	75.6	510	1	FOLJTM	gag polyprotein -
38	34	75.6	521	1	FOLJST	gag polyprotein -
39	34	75.6	521	1	FOLJCA	gag polyprotein -
40	34	75.6	521	2	S08435	gag polyprotein -
41	34	75.6	521	2	S12152	gag polyprotein -
42	34	75.6	521	2	S53091	gag polyprotein -
43	34	75.6	522	1	FOLJG2	gag polyprotein -
44	34	75.6	522	1	FOLJG6	gag polyprotein -
45	34	75.6	557	2	H70725	hypothetical prote

#### RESULT 1

S24478 gag polyprotein - human immunodeficiency virus type 1 (fragment)

N:Contains: p24 protein; p25 protein; p7 protein HIV-1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 20-Feb-1995 #sequence\_revision 26-Jul-1996 #text\_change 26-Aug-1999

C:Accession: S24478; S24484

R:Salminen, M.

Submitted to the EMBL Data Library, October 1991

A:Reference number: S24471

A:Accession: S24478

A:Molecule type: DNA

A:Residues: 1-75 <SAL>

A:Cross-references: EMBL:Z11148; NID:960078; PIDN:CAAT7499.1; PID:939876

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: polyprotein

Query Match

Best Local Similarity 88.9%; Score 40; DB 2; Length 75;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10

DB 3 KARVLAEM 11

#### RESULT 2

S24475 gag polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S24475; S24488

R:Salminen, M.

Submitted to the EMBL Data Library, October 1991

A:Reference number: S24471

A:Accession: S24475

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 <SAL>

A:Cross-references: EMBL:Z11142; NID:960101; PIDN:CAAT77499.1; PID:960102

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: polyprotein

Query Match

Best Local Similarity 88.9%; Score 40; DB 2; Length 75;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10

DB 2 KARVLAEM 10

## RESULT 3

S24474

gag polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C&gt;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S24474; S24487

R:Salminen, M.

submitted to the EMBL Data Library, October 1991

A:Reference number: S24474

A:Accession: S24474

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 &lt;SAL&gt;

A:Cross-references: EMBL:Z11138; NID:960095; PIDN:CAA77489.1; PID:960096

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: polyprotein

Query Match

Best Local Similarity 100.0%; Score 40; DB 2; Length 75; Pred. No. 0.5; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10

DB 2 KARVLAEM 10

## RESULT 4

S24476

gag polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C&gt;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S24476; S24489

R:Salminen, M.

submitted to the EMBL Data Library, October 1991

A:Reference number: S24471

A:Accession: S24476

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-76 &lt;SAL&gt;

A:Cross-references: EMBL:Z11143; NID:960103; PIDN:CAA77494.1; PID:960104

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: polyprotein

Query Match

Best Local Similarity 100.0%; Score 40; DB 2; Length 76; Pred. No. 0.51; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10

DB 2 KARVLAEM 10

## RESULT 5

FOVWVL

gag polyprotein - human immunodeficiency virus type 1 (isolate LV)

N:Alternate names: assemblin; core polyprotein; gag precursor

N:Contains: capsid antigen core protein p24CA; core protein p1; core protein p6; matrix

C:Species: human immunodeficiency virus type 1, HIV-1

C&gt;Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 05-Dec-1998

C:Accession: A03948

R:Musling, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.

Nature 313, 450-458, 1985

A:Title: Nucleic acid structure and expression of the human AIDS/Lymphadenopathy retrov

A:Reference number: A93355; MUID:8511157; PMID:2982104

A:Accession: A03948

A:Molecule type: DNA

A:Residues: 1-478 &lt;MOE&gt;

C:Genetics:

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: AIDS; blocked amino end; core protein; immunodeficiency; 11popenin; myrist

F:2-478/Product: gag precursor (assemblin) #status predicted &lt;GAG&gt;

F:2-132/Product: matrix antigen core protein p17MA #status predicted &lt;p17&gt;

F:20-32/Region: nuclear location signal

F:110-114/Region: nuclear location signal

F:133-363/Product: capsid antigen core protein p24CA #status predicted &lt;p24&gt;

F:364-377/Product: core protein p2 #status predicted &lt;CP2&gt;

F:378-433/Product: nucleocapsid core protein p7NC #status predicted &lt;CP7&gt;

F:392-405/Region: zinc finger CCHC motif

F:413-426/Region: zinc finger CCHC motif

F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

F:392,395,400,405/Binding site: zinc (Cys, Cys, His, Cys) #status experimental

F:413,416,421,426/Binding site: zinc (Cys, Cys, His, Cys) #status experimental

Query Match

Best Local Similarity 100.0%; Score 40; DB 1; Length 478; Pred. No. 3.2; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10

DB 356 KARVLAEM 367

## RESULT 6

FOJUND

gag polyprotein - human immunodeficiency virus type 1 (isolate NDK)

N:Alternate names: core polyprotein

N:Contains: core protein p15; core protein p17; core protein p24

C:Species: human immunodeficiency virus type 1, HIV-1

C&gt;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999

C:Accession: J00065

R:Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe,

Gene 81, 275-284, 1989

A:Title: Nucleotide sequence of HIV-1-NK: a highly cytopathic strain of the human imm

A:Reference number: J00065; MUID:90034200; PMID:2806917

A:Accession: J00065

A:Molecule type: DNA

A:Residues: 1-497 &lt;SPI&gt;

A:Cross-references: GB:M27323; NID:9328154; PIDN:AAA44868.1; PID:9328157

C:Genetics:

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: AIDS; core protein; immunodeficiency; polyprotein

F:1-129/Product: core protein p17 #status predicted &lt;C17&gt;

F:130-389/Product: core protein p24 #status predicted &lt;C24&gt;

F:390-497/Product: core protein p15 #status predicted &lt;C15&gt;

Query Match

Best Local Similarity 100.0%; Score 40; DB 1; Length 497; Pred. No. 3.4; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10

DB 356 KARVLAEM 364

## RESULT 7

T09436

gag polyprotein - human immunodeficiency virus type 1 (strain JRF1)

C:Species: human immunodeficiency virus type 1, HIV-1

C&gt;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 24-Nov-1999

C:Accession: T09436

R:Pang, S.; Vinters, H.V.; Akashl, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namer

submitted to the EMBL Data Library, July 1996

A:Reference number: Z16673

A:Accession: T09436

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-498 &lt;PAN&gt;

A:Cross-references: EMBL:U63632; NID:g1465777; PID:g1465778

C:Genetics:

C:Superfamily: AIDS-related virus gag polyprotein

Query Match 88.9%; Score 40; DB 2; Length 498;

Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KARVLAEM 10  
Db 359 KARVLAEM 365

## RESULT 8

FOYMLV

gag polyprotein - human immunodeficiency virus type 1 (isolate LAV-1a)

N:Alternate names: assemblin; core polyprotein; gag precursor

N:Contains: capsid antigen core protein p24CA; core protein p1; core protein p6; matrix

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 05-Dec-1998

C:Accession: A03946

R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.

Cell 40, 9-17, 1985

A:Title: Nucleotide sequence of the AIDS virus, LAV.

A:Reference number: A90866; PMID:8509333; PMID:2981635

A:Accession: A03946

A:Molecule type: DNA

A:Residues: 1-500 <MAI>

C:Genetics:

A:Gene: gag

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: AIDS; blocked amino end; core protein; immunodeficiency; lipoprotein; myrist

F:2-500/Product: gag precursor (assemblin) #status predicted <GAG>

F:2-133/Product: matrix antigen core protein p17MA #status predicted <P17>

F:20-33/Region: nuclear location signal

F:110-114/Region: nuclear location signal

F:133-363/Product: capsid antigen core protein p24CA #status predicted <P24>

F:364-377/Product: core protein p2 #status predicted <CP2>

F:378-432/Product: nucleocapsid core protein p7NC #status predicted <CP7>

F:392-405/Region: zinc finger CCHC motif

F:413-426/Region: zinc finger CCHC motif

F:433-448/Product: core protein p1 #status predicted <CP1>

F:449-500/Product: core protein p6 #status predicted <CP6>

F:2/Modified site: myristylated amino end (gly) (in mature form) #status predicted

F:392,395,400,405/Binding site: zinc (Cys, Cys, His, Cys) #status experimental

F:413,416,421,426/Binding site: zinc (Cys, Cys, His, Cys) #status experimental

Query Match 88.9%; Score 40; DB 1; Length 500;

Best Local Similarity 100.0%; Pred. No. 3.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
Db 359 KARVLAEM 367

## RESULT 9

gag polyprotein - human immunodeficiency virus type 1 (strain YU-2)

N:Alternate names: core polyprotein

N:Contains: core protein p15; core protein p17; core protein p24

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 31-Jan-1997

C:Accession: A44001

R:Li, Y.; Hul, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.

J. Virol. 66, 6587-6600, 1992

A:Title: Complete nucleotide sequence, genome organization, and biological properties of

A:Reference number: A44001; PMID:93021387; PMID:1404605

A:Accession: A44001

A:Molecule type: DNA

A:Residues: 1-500 <LIV>

A:Cross-references: GB:M93258

C:Genetics:

A:Gene: gag

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: AIDS; core protein; immunodeficiency; polyprotein

F:1-133/Product: core protein p17 #status predicted <P17>  
F:133-391/Product: core protein p24 #status predicted <P24>  
F:392-500/Product: core protein p15 #status predicted <P15>

Query Match 88.9%; Score 40; DB 1; Length 500;

Best Local Similarity 100.0%; Pred. No. 3.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
Db 359 KARVLAEM 367

## RESULT 10

FOYHHA

gag polyprotein - human immunodeficiency virus type 1 (isolate CDC-451)

N:Alternate names: core polyprotein

N:Contains: core protein p15; core protein p17; core protein p24

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 16-Jul-1999

C:Accession: A25523

R:Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devay

Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986

A:Title: Molecular cloning and primary nucleotide sequence analysis of a distinct hu

A:Reference number: A94136; PMID:87041461; PMID:3490666

A:Accession: A25523

A:Molecule type: DNA

A:Residues: 1-500 <DES>

A:Cross-references: GB:M13136; NID:9326459; PIDN:AAA44306.1; PID:9326462

C:Genetics:

A:Gene: gag

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: core protein; polyprotein

F:1-133/Product: core protein p17 #status predicted <P17>

F:133-391/Product: core protein p24 #status predicted <P24>

F:392-500/Product: core protein p15 #status predicted <P15>

Query Match 88.9%; Score 40; DB 1; Length 500;

Best Local Similarity 100.0%; Pred. No. 3.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
Db 359 KARVLAEM 367

## RESULT 11

S33979

gag polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999

C:Accession: S33979

R:Carlini, F.

submitted to the EMBL Data Library, November 1991

A:Reference number: S33979

A:Accession: S33979

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-500 <CAR>

A:Cross-references: EMBL:Z11530; NID:960192; PIDN:CAA77621.1; PID:960193

C:Superfamily: AIDS-related virus gag polyprotein

Query Match 88.9%; Score 40; DB 2; Length 500;

Best Local Similarity 100.0%; Pred. No. 3.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
Db 359 KARVLAEM 367

## RESULT 12

S54377

gag polyprotein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 26-Aug-1999  
 C:Accession: S54377  
 R:Theodore, T.; Buckler-White, A.J.  
 submitted to the EMBL Data Library, July 1989  
 A:Reference number: S54377  
 A:Accession: S54377  
 A:Status: preliminary  
 A:Molecule type: genomic RNA  
 A:Residues: 1-501 <THR>  
 A:Cross-references: EMBL:M2639; NID:g329377; PIDN:AAA5365.1; PID:g329380  
 C:Superfamily: AIDS-related virus gag polyprotein  
 C:Keywords: polyprotein

Query Match 88.9%; Score 40; DB 2; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 |||||  
 DB 360 KARVLAEM 368

RESULT 13  
 FOVMA2  
 gag polyprotein - human immunodeficiency virus type 1 (isolate ARV-2)

N:Alternate names: core polyprotein  
 N:Contains: core protein p15; core protein p17; core protein p24  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A>Note: host Homo sapiens (man)  
 C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
 C:Accession: A03947

R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-St  
 Science 227, 484-492, 1985  
 A>Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).  
 A:Reference number: A04003; MUID:85090453; PMID:2578227  
 A:Accession: A03947  
 A:Molecule type: DNA

A:Residues: 1-502 <SAN>  
 A:Cross-references: GB:K02007; NID:g328658; PIDN:AA59875.1; PID:g328661  
 C:Comment: Cleavage sites that yield the mature core proteins remain to be determined.  
 C:Genetics:

A:Gene: gag  
 C:Superfamily: AIDS-related virus gag polyprotein  
 C:Keywords: AIDS; core protein; immunodeficiency; polyprotein  
 F:1-134/Product: core protein p17 #status predicted <p17>  
 F:135-393/Product: core protein p24 #status predicted <p24>  
 F:394-502/Product: core protein p15 #status predicted <p15>

Query Match 88.9%; Score 40; DB 1; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 |||||  
 DB 361 KARVLAEM 369

RESULT 14  
 T01667

gag polyprotein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 21-Jul-2000  
 C:Accession: T01667  
 R:Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.  
 Cell 46, 63-74, 1986

A>Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two iso  
 A:Reference number: Z14389; MUID:86245056; PMID:2424612  
 A:Accession: T01667

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-505 <ALI>

A:Cross-references: EMBL:K03456; NID:g60228; PIDN:CAA28011.1; PID:g60229  
 C:Superfamily: AIDS-related virus gag polyprotein

Query Match 88.9%; Score 40; DB 2; Length 505;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 |||||  
 DB 365 KARVLAEM 373

RESULT 15  
 A38068  
 gag polyprotein - human immunodeficiency virus type 1 (strain MN)

N:Alternate names: core polyprotein  
 N:Contains: core protein p1; core protein p17; core protein p2; core protein p24; cor  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A>Note: host Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 26-Feb-1999  
 C:Accession: A38068

R:Henderson, L.E.; Bowers, M.A.; Sowder II, R.C.; Serabyn, S.A.; Johnson, D.G.; Bess  
 J. Virol. 66, 1856-1865, 1992  
 A>Title: Gag proteins of the highly replicative MN strain of human immunodeficiency v  
 A:Reference number: A38068; MUID:92194415; PMID:1548743  
 A:Accession: A38068

A:Molecule type: protein  
 A:Residues: 1-506 <HEM>  
 C:Genetics:

A:Gene: gag  
 C:Superfamily: AIDS-related virus gag polyprotein  
 C:Keywords: AIDS; core protein; immunodeficiency; polyprotein  
 F:1-134/Product: core protein p17 #status experimental <p17>  
 F:135-365/Product: core protein p24 #status experimental <p24>  
 F:366-379/Product: core protein p2 #status experimental <pp2>  
 F:380-434/Product: core protein p7 #status experimental <pp7>  
 F:435-450/Product: core protein p1 #status experimental <pp1>  
 F:451-506/Product: core protein p6 #status experimental <pp6>

Query Match 88.9%; Score 40; DB 1; Length 506;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 |||||  
 DB 361 KARVLAEM 369

Search completed: March 13, 2003, 18:52:59  
 Job time : 9.81818 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:17 ; Search time 9.81818 Seconds  
(without alignments)  
97.915 Million cell updates/sec

Title: US-10-050-200-5

Perfect score: 44

Sequence: 1 EKARVLAEEA 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

PIR-73:\*  
1: pirl:\*  
2: pirl:\*  
3: pirl:\*  
4: pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	86.4	122	2	S72820 hypothetical prote
2	38	86.4	263	2	S72978 hypothetical prote
3	38	86.4	281	2	G86956 conserved hypothet
4	38	86.4	557	2	H70725 hypothetical prote
5	36	81.8	141	2	B97591 50S ribosomal prot
6	36	81.8	141	2	AH2812 50S ribosomal prot
7	36	81.8	298	2	G81895 Probable acetylglu
8	36	81.8	298	2	E81124 acetylglutamate ki
9	35	79.5	75	2	S24475 gag polypotein -
10	35	79.5	75	2	S24475 gag polypotein -
11	35	79.5	76	2	S24474 gag polypotein -
12	35	79.5	423	2	B72528 gag polypotein -
13	35	79.5	478	1	FOVWVL Probable nucleolar
14	35	79.5	497	1	FOVJND gag polypotein -
15	35	79.5	498	1	FOVWVY gag polypotein -
16	35	79.5	500	1	FOVWVY gag polypotein -
17	35	79.5	500	1	FOVWVY gag polypotein -
18	35	79.5	500	1	FOVWVY gag polypotein -
19	35	79.5	500	1	FOVWVY gag polypotein -
20	35	79.5	500	1	FOVWVY gag polypotein -
21	35	79.5	501	2	S54377 gag polypotein -
22	35	79.5	502	1	FOVWVY gag polypotein -
23	35	79.5	505	2	T01667 gag polypotein -
24	35	79.5	506	1	A38068 gag polypotein -
25	35	79.5	508	1	FOVJSI gag polypotein -
26	35	79.5	512	1	FOVWVY gag polypotein -
27	35	79.5	1021	2	H75423 hypothetical prote
28	34	77.3	399	2	E86289 t16N1.7 protein -
29	34	77.3	785	2	B84038 DNA mismatch repai

30	34	77.3	808	2	F70720 probable transloca
31	34	77.3	862	2	A05028 ipoc protein homol
32	34	77.3	1163	2	S07137 DNA-directed RNA p
33	34	77.3	1224	2	T07446 DNA-directed RNA p
34	34	77.3	1361	2	A29959 DNA-directed RNA p
35	34	77.3	1513	1	RNRZC2 DNA-directed RNA p
36	34	77.3	1527	1	RNZMB2 DNA-directed RNA p
37	33	75.0	126	2	C95270 hypothetical prote
38	33	75.0	236	2	H86706 hypothetical prote
39	33	75.0	260	2	C69130 shikimate 5-dehydr
40	33	75.0	386	2	S47163 DNA-directed RNA p
41	33	75.0	393	2	S25565 DNA-directed RNA p
42	33	75.0	397	2	C75182 DNA-directed RNA p
43	33	75.0	397	2	G71031 probable DNA-direc
44	33	75.0	409	2	E72765 probable phosphos
45	33	75.0	516	2	A31270 radial spoke prote

## ALIGNMENTS

## RESULT 1

S72820

hypothetical protein B1620.C3.232 - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 23-Mar-2001

C:Accession: S72820

R:Smith, D.R.; Robinson, K.

submitted to the EMBL Data Library, November 1993

A:Description: Mycobacterium leprae cosmid B1620.

A:Reference number: S72584

A:Accession: S72820

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-122 <SMI>

A:Cross-references: EMBL:U00015; NID:g466931; PIDN:AAC43229.1; PID:g466941

C:Genetics:

A:Start codon: GTG

Query Match Best Local Similarity 86.4%; Score 38; DB 2; Length 122;

Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEEA 10

DB 59 ERARVVAEEA 68

## RESULT 2

S72978

hypothetical protein B229.C1.175 - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 23-Mar-2001

C:Accession: S72978

R:Smith, D.R.; Robinson, K.

submitted to the EMBL Data Library, November 1993

A:Description: Mycobacterium leprae cosmid B229.

A:Reference number: S72588

A:Accession: S72978

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-263 <SMI>

A:Cross-references: EMBL:U00020

C:Genetics:

A:Start codon: GTG

Query Match Best Local Similarity 86.4%; Score 38; DB 2; Length 263;

Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEEA 10

DB 218 ERARVVAEEA 227

RESULT 3  
G86956  
conserved hypothetical protein ML0383 [Imported] - Mycobacterium lepre  
C:Species: Mycobacterium lepre  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: G86956  
R: Cole, S.T.; Elgimeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hc  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A: Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
A: Title: Massive gene decay in the leprosy bacillus.  
A: Reference number: AB6909; MUID:21128732; PMID:11234002  
A: Accession: G86956  
A: Status: Preliminary  
A: Molecule type: DNA  
A: Residues: 1281 <STO>  
A: Cross-references: GB:AL450380; NID:g13092660; PIDN:CAC29891.1; GSPDB:GN00147  
C: Genetics:  
A: Gene: ML0383

Query Match      86.4%;      Score 38; DB 2;      Length 281;  
Best Local Similarity      80.0%;      Pred. No. 5.1;  
Matches      8;      Conservative      2;      Mismatches      0;      Indels      0;      Gaps      0;

OY      1 EKARVLAEMA 10  
      |||||  
Db      218 ERARVVAEMA 227

RESULT 4  
H70725  
Hypothetical protein RV2585c - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: H70725  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A: Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID:98295987; PMID:9634230  
A: Accession: H70725  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-557 <COI>  
A: Cross-references: GB:Z77724; GB:AL123456; NID:g3261620; PIDN:CAB01258.1; PID:e256146;  
A: Experimental source: strain H37RV  
C: Genetics:  
A: Gene: RV2585c

Query Match      86.4%;      Score 38; DB 2;      Length 557;  
Best Local Similarity      80.0%;      Pred. No. 10;  
Matches      8;      Conservative      2;      Mismatches      0;      Indels      0;      Gaps      0;

OY      1 EKARVLAEMA 10  
      |||||  
Db      500 ERARVLAEMA 509

RESULT 5  
B97591  
50S ribosomal protein L17 nma0102 [Imported] - Agrobacterium tumefaciens (strain C58, Ce  
C:Species: Agrobacterium tumefaciens  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: B97591  
R: Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Matkeltz, B.;  
Science 294, 2333-2328, 2001  
A: Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A: Reference number: A97359; PMID:11743194  
A: Accession: B97591  
A: Status: Preliminary  
A: Molecule type: DNA  
A: Residues: 1-141 <KUR>  
A: Cross-references: GB:AE007869; PIDN:AAK87683.1; PID:g15157042; GSPDB:GN00169  
C: Genetics:  
A: Gene: AGR\_C-3516  
A: Map position: circular chromosome  
C: Superfamily: Escherichia coli ribosomal protein L17

Query Match      81.8%;      Score 36; DB 2;      Length 141;  
Best Local Similarity      80.0%;      Pred. No. 6.5;  
Matches      8;      Conservative      1;      Mismatches      1;      Indels      0;      Gaps      0;

OY      1 EKARVLAEMA 10  
      |||||  
Db      127 DKARVVAEMA 136

RESULT 6  
AH2812  
50S ribosomal protein L17 [Imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
C:Accession: AH2812  
R: Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; Mocl  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A: Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.  
A: Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A: Reference number: AB2577; PMID:11743193  
A: Accession: AH2812  
A: Status: Preliminary  
A: Molecule type: DNA  
A: Residues: 1-141 <KUR>  
A: Cross-references: GB:AE008688; PIDN:AL42918.1; PID:g17740374; GSPDB:GN00186  
A: Experimental source: strain C58 (Dupont)  
C: Genetics:  
A: Gene: rplQ  
A: Map position: circular chromosome  
C: Superfamily: Escherichia coli ribosomal protein L17

Query Match      81.8%;      Score 36; DB 2;      Length 141;  
Best Local Similarity      80.0%;      Pred. No. 6.5;  
Matches      8;      Conservative      1;      Mismatches      1;      Indels      0;      Gaps      0;

OY      1 EKARVLAEMA 10  
      |||||  
Db      127 DKARVVAEMA 136

RESULT 7  
G81895  
Probable acetylglutamate kinase (EC 2.7.2.8) NMA1275 [Imported] - Neisseria meningitidis  
C:Species: Neisseria meningitidis  
C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: G81895  
R: Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo  
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream  
Nature 404, 502-506, 2000  
A: Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis z2491  
A: Reference number: AB1775; MUID:20222556; PMID:10761919  
A: Accession: G81895  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-298 <PAR>  
A: Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84528.1; PID:g737  
A: Experimental source: serogroup A, strain z2491  
C: Genetics:  
A: Gene: argB; NMA1275

C:Superfamily: acetylglutamate kinase  
C:Keywords: phosphotransferase

Query Match 81.8%; Score 36; DB 2; Length 298;  
Best Local Similarity 77.8%; Pred. No. 14;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEA 9  
:|||||  
Db 11 DKARVLAEA 19

RESULT 8  
E81124  
acetylglutamate kinase NMB1074 [imported] - Neisseria meningitidis (strain MC58 serogroup B)

C:Species: Neisseria meningitidis  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: E81124

R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Vignatelli, B.  
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: E81124

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-298 <TEXT>

A:Cross-references: GB:AE002458; GB:AE002098; NID:g7226311; PIDN:AAF41469.1; PID:g7226311

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB1074

C:Superfamily: acetylglutamate kinase

Query Match 81.8%; Score 36; DB 2; Length 298;  
Best Local Similarity 77.8%; Pred. No. 14;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEA 9  
:|||||  
Db 11 DKARVLAEA 19

RESULT 9

S24478  
gag polyprotein - human immunodeficiency virus type 1 (fragment)

N:Contains: p24 protein; p25 protein; p7 protein

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 20-Feb-1995 #sequence\_revision 26-Jul-1996 #text\_change 26-Aug-1999

C:Accession: S24478; S24484

R:Salminen, M.

submitted to the EMBL Data Library, October 1991

A:Reference number: S24471

A:Accession: S24478

A:Molecule type: DNA

A:Residues: 1-75 <SAL>

A:Cross-references: EMBL:Z11148; NID:g60078; PIDN:CAA77499.1; PID:g939876

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: polyprotein

Query Match 79.5%; Score 35; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
:|||||  
Db 3 KARVLAEA 10

RESULT 10

S24475  
gag polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S24475; S24488

R:Salminen, M.

submitted to the EMBL Data Library, October 1991

A:Reference number: S24471

A:Accession: S24475

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 <SAL>

A:Cross-references: EMBL:Z11142; NID:g60101; PIDN:CAA77493.1; PID:g60102

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: polyprotein

Query Match 79.5%; Score 35; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
:|||||  
Db 2 KARVLAEA 9

RESULT 11

S24474  
gag polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S24474; S24487

R:Salminen, M.

submitted to the EMBL Data Library, October 1991

A:Reference number: S24471

A:Accession: S24474

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 <SAL>

A:Cross-references: EMBL:Z11138; NID:g60095; PIDN:CAA77489.1; PID:g60096

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: polyprotein

Query Match 79.5%; Score 35; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
:|||||  
Db 2 KARVLAEA 9

RESULT 12

S24476  
gag polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S24476; S24489

R:Salminen, M.

submitted to the EMBL Data Library, October 1991

A:Reference number: S24471

A:Accession: S24476

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-76 <SAL>

A:Cross-references: EMBL:Z11143; NID:g60103; PIDN:CAA77494.1; PID:g60104

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: polyprotein

Query Match 79.5%; Score 35; DB 2; Length 76;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
:|||||  
Db 2 KARVLAEA 9

## RESULT 13

B72528

probable nucleolar protein NOP5 APE2199 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: B72528

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jinno, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A&gt;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: B72528

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-423 &lt;KAN&gt;

A:Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BAAB1210.1; PID:95105898

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2199

C:Superfamily: garden pea SAR DNA-binding protein

Query Match

Best Local Similarity 79.5%; Score 35; DB 2; Length 423;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EKARVLAEA 10

DB 218 EKARVLAEA 227

## RESULT 14

FOVWYL

gag polyprotein - human immunodeficiency virus type 1 (isolate LV)

N:Alternate names: assemblin; core polyprotein; gag precursor

N:Contains: capsid antigen core protein p24CA; core protein p1; core protein p6; matrix

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 05-Dec-1998

C:Accession: A03948

R:Musning, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.

Nature 313, 450-458, 1985

A&gt;Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovi

A:Reference number: A93355; MUID:85111157; PMID:2982104

A:Accession: A03948

A:Molecule type: DNA

A:Residues: 1-478 &lt;MUR&gt;

C:Genetics:

A:Gene: gag

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: AIDS; blocked amino end; core protein; immunodeficiency; lipoprotein; myrist

F:2-478/Product: gag precursor (assemblin) #status predicted &lt;GAG&gt;

F:2-132/Product: matrix antigen core protein p17MA #status predicted &lt;P17&gt;

F:20-32/Region: nuclear location signal

F:110-114/Region: nuclear location signal

F:133-363/Product: capsid antigen core protein p24CA #status predicted &lt;P24&gt;

F:364-377/Product: core protein p2 #status predicted &lt;CP2&gt;

F:378-432/Product: nucleocapsid core protein p7NC #status predicted &lt;CP7&gt;

F:393-405/Region: zinc finger CCHC motif

F:413-426/Region: zinc finger CCHC motif

F:2/Modified site: myristylated amino end (gly) (in mature form) #status predicted

F:392,395,400,405/Binding site: zinc (Cys, Cys, His, Cys) #status experimental

F:413,416,421,426/Binding site: zinc (Cys, Cys, His, Cys) #status experimental

Query Match

Best Local Similarity 79.5%; Score 35; DB 1; Length 478;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9

DB 359 KARVLAEA 366

## RESULT 15

FOLJND

gag polyprotein - human immunodeficiency virus type 1 (isolate NDK)

N:Alternate names: core polyprotein

N:Contains: core protein p15; core protein p17; core protein p24

C:Species: human immunodeficiency virus type 1, HIV-1

A&gt;Note: host Homo sapiens (man)

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999

C:Accession: J00065

R:Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe,

Gene 81, 275-284, 1989

A&gt;Title: Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human imm

A:Reference number: J00065; MUID:90034200; PMID:2806917

A:Accession: J00065

A:Molecule type: DNA

A:Residues: 1-497 &lt;SP1&gt;

A:Cross-references: GB:M27323; NID:9328154; PIDN:AAA44868.1; PID:9328157

C:Genetics:

A:Gene: gag

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: AIDS; core protein; immunodeficiency; polyprotein

F:1-129/Product: core protein p17 #status predicted &lt;C17&gt;

F:130-389/Product: core protein p24 #status predicted &lt;C24&gt;

F:390-497/Product: core protein p15 #status predicted &lt;C15&gt;

Query Match

Best Local Similarity 79.5%; Score 35; DB 1; Length 497;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9

DB 356 KARVLAEA 363

Search completed: March 13, 2003, 18:52:59  
Job time : 11.8182 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:45:22 ; Search time 20.7273 Seconds  
(without alignments)  
99.409 Million cell updates/sec

Title: US-10-050-200-5  
Perfect score: 44  
Sequence: 1 EKARVLAEAA 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriaph:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	90.9	152	16	0921E1
2	38	86.4	122	2	049726
3	38	86.4	264	2	049858
4	38	86.4	281	16	09CCV3
5	38	86.4	396	4	096K21
6	37	84.1	637	17	09HG28
7	36	81.8	141	16	08UE41
8	36	81.8	262	16	092NA6
9	35	79.5	46	15	09J122
10	35	79.5	54	15	09J181
11	35	79.5	55	15	09J140
12	35	79.5	56	15	09J117
13	35	79.5	58	15	09J144
14	35	79.5	58	15	09J138
15	35	79.5	61	15	09J198
16	35	79.5	61	15	09J197

17	35	79.5	61	15	09J168
18	35	79.5	62	15	09J132
19	35	79.5	62	15	09J612
20	35	79.5	63	15	09J177
21	35	79.5	63	15	09J126
22	35	79.5	63	15	09J123
23	35	79.5	64	15	09J189
24	35	79.5	64	15	09J184
25	35	79.5	64	15	09J182
26	35	79.5	64	15	09J167
27	35	79.5	64	15	09J159
28	35	79.5	64	15	09J155
29	35	79.5	65	15	09J195
30	35	79.5	65	15	09J174
31	35	79.5	65	15	09J171
32	35	79.5	65	15	09J160
33	35	79.5	65	15	09J158
34	35	79.5	65	15	09J147
35	35	79.5	65	15	09J139
36	35	79.5	65	15	09J137
37	35	79.5	65	15	09J136
38	35	79.5	65	15	09J120
39	35	79.5	65	15	09J115
40	35	79.5	66	15	09J186
41	35	79.5	66	15	09J175
42	35	79.5	66	15	09J173
43	35	79.5	66	15	09J127
44	35	79.5	66	15	09J125
45	35	79.5	66	15	09J597

## ALIGNMENTS

## RESULT 1

0921E1 PRELIMINARY: PRT: 152 AA.  
AC 0921E1;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Putative transposase number 1 for insertion sequence ISRM28.  
GN TRM28.1 OR R03124 OR SMC03278.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_Taxid:382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-1021;  
RX MEDLINE-21396507; PubMed-11481430;  
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J., Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S., Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D., Pohl T., Portetalle D., Puehler A., Purnelle B., Rampeger U., Renaud C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;  
RA "Analysis of the chromosome sequence of the legume symbiont Sinorhizobium meliloti strain 1021.";  
RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
DR EMBL: AL591793; CAC47703.1;  
DR InterPro: IPR002514; Transposase\_8.  
DR Pfam: PF01527; Transposase\_8; 1.  
KW Complete proteome.  
SQ SEQUENCE 152 AA; 16544 MW; F518578357FF0C5C CRC64;

Query Match 90.9%; Score 40; DB 16; Length 152;  
Best Local Similarity 80.0%; Pred. No. 3.4;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEAA 10  
DB 41 EKARVLAEAA 50

```

RESULT 2
ID 049726 PRELIMINARY; PRT; 122 AA.
AC 049726:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-FEB-1997 (TREMBLrel. 02, Last annotation update)
DE Hypothetical 12.9 kDa protein B1620_C3_232 precursor.
GN B1620_C3_232.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriales; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO M.TUBERCULOSIS MYCY8.14.
DR EMBL; 000015; AAC33229.1; -.
DE Hypothetical protein; Signal.
FT SIGNAL 1 20
FT CHAIN 1 20 HYPOTHETICAL PROTEIN B1620_C3_232.
SQ SEQUENCE 122 AA; 12987 MW; 4C7D322658C06A99 CRC64;

Query Match
Best Local Similarity 86.4%; Score 38; DB 2; Length 122;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKARVLAEEA 10
DB 59 ERARVVAEEA 68
I:|||||
|:|||||

RESULT 3
ID 049858 PRELIMINARY; PRT; 264 AA.
AC 049858:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-FEB-1997 (TREMBLrel. 02, Last annotation update)
DE Hypothetical 27.4 kDa protein B229_C1_175 precursor.
GN B229_C1_175.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriales; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO M.TUBERCULOSIS MYCY8.14.
DR EMBL; 000020; AAA17292.1; -.
DE Hypothetical protein; Signal.
FT SIGNAL 1 23
FT CHAIN 1 23 HYPOTHETICAL PROTEIN B229_C1_175.
SQ SEQUENCE 264 AA; 27430 MW; 7BBA3900F891IDF4 CRC64;

Query Match
Best Local Similarity 86.4%; Score 38; DB 2; Length 264;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKARVLAEEA 10
DB 218 ERARVVAEEA 227
I:|||||
|:|||||

RESULT 4
ID 09CCV3 PRELIMINARY; PRT; 281 AA.
AC 09CCV3:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

```

```

DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Hypothetical protein ML0383.
GN ML0383.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriales; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT *Massive gene decay in the leprosy bacillus.*;
RL Nature 409:1007-1011(2001).
DR EMBL; AL583918; CAC29891.1; -.
DE Leprosy; ML0383;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 281 AA; 29333 MW; FE0A1BDCFD69803F CRC64;

Query Match
Best Local Similarity 86.4%; Score 38; DB 16; Length 281;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKARVLAEEA 10
DB 218 ERARVVAEEA 227
I:|||||
|:|||||

RESULT 5
ID 096K21 PRELIMINARY; PRT; 396 AA.
AC 096K21:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CDNA FLJ14840 f1s, clone OVARC1001916 (Hypothetical 43.9 kDa
DE protein).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagaitsuma M., Hosoliri T., Raku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishi S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
RA Ninomiya K., Iwayanagi T.;
RT *NEO human cDNA sequencing project.*;
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=RETINA;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK027746; BAB55338.1; -.
DR EMBL; BC015738; AAH15738.1; -.
DR InterPro: IPR000564; 2FE2S, ferredoxin.
DR InterPro: IPR000345; CyfC_heme_bind.
DR InterPro: IPR000306; ZnF_FYVE.
DR Pfam: PF01363; FYVE; 1.
DR PROSITE: PS00197; 2FE2S_FERREDOXIN; UNKNOWN_1.

```

```
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 396 AA; 43946 MW; FA4ECBA371A5A977 CRC64;

Query Match
Best Local Similarity 86.4%; Score 38; DB 4; Length 396;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEEA 10
DB 229 EKSRVLAEEA 238

RESULT 6
O9HGZ8 PRELIMINARY; PRT; 637 AA.
ID O9HGZ8;
AC O9HGZ8;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical protein PF1746 (Hypothetical 72.7 kDa protein).
GN PF1746.
OS Pyrococcus furiosus, and
OC Thermococcus litoralis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261, 2265;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=P.furiosus, and T.litoralis;
RX MEDLINE=20566786; PubMed=11115105;
RA Druggiero J., Dunn D., Maeder D.L., Holley-Shanks R., Chataud J.,
RA Horlacher R., Robb F.T., Boos W., Weiss R.B.;
RT "Evidence of recent lateral gene transfer among hyperthermophilic
RT archaea."
RL MOL. Microbiol. 38:684-693(2000).
DR EMBL: AE010273; AAL81870.1; -
DR EMBL: AF307053; AAG45395.1; -
DR EMBL: AF307052; AAG45379.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 637 AA; 72701 MW; EBF42F42FB5A428 CRC64;

Query Match
Best Local Similarity 84.1%; Score 37; DB 17; Length 637;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEEA 9
DB 531 EKARVLAEEA 539

RESULT 7
O8UE41 PRELIMINARY; PRT; 141 AA.
ID O8UE41;
AC O8UE41;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE 50S ribosomal protein L17.
GN RPLQ OR ATU1922 OR AGR.C 3516.
OS Aerobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Bpp A., Liu F.,
RA Houlam C., Allinger M., Doughty D., Scott C., Iappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cleto C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL: AE009146; AAL42918.1; -
DR EMBL: AE008111; AAK87683.1; -
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 141 AA; 15395 MW; CD37B5DAE2D79D35 CRC64;

Query Match
Best Local Similarity 81.8%; Score 36; DB 16; Length 141;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EKARVLAEEA 10
DB 127 DKARVLAEEA 136

RESULT 8
O92NA6 PRELIMINARY; PRT; 262 AA.
ID O92NA6;
AC O92NA6;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Hypothetical protein R02307.
GN R02307 OR SMC01556.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaire V., Masny D.,
RA Pohl T., Portetle D., Puehler A., Punelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL: AL591790; CAC46886.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 262 AA; 27103 MW; 986D78DEDDDE79F2 CRC64;

Query Match
Best Local Similarity 81.8%; Score 36; DB 16; Length 262;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEEA 10
```

Db 197 KAKVLAEEA 205

11:111111

## RESULT 9

ID 09J122 PRELIMINARY: PRT: 46 AA.  
AC 09J122;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE Gag protein (Fragment).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FIN92168;  
RA Lillsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
RA Leinikki P., Salminen M.O.;  
RT \*Analysis of HIV-1 genetic subtypes in Finland reveals good  
RT correlation between molecular and epidemiological data.\*;  
RL Scand. J. Infect. Dis. 0:0-0(2000).  
DR EMBL: AF219351; AAF30257.1; -  
FT NON\_TER 1 1  
SQ SEQUENCE 46 AA; 4946 MW; 7F85C2C145AF1677 CRC64;

Query Match 79.5%; Score 35; DB 15; Length 46;  
Best Local Similarity 100.0%; Pred. No. 9.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KAKVLAEEA 9  
Db 6 KAKVLAEEA 13

## RESULT 10

ID 09J181 PRELIMINARY: PRT: 54 AA.  
AC 09J181;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Gag protein (Fragment).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FIN9140;  
RA Lillsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
RA Leinikki P., Salminen M.O.;  
RT \*Analysis of HIV-1 genetic subtypes in Finland reveals good  
RT correlation between molecular and epidemiological data.\*;  
RL Scand. J. Infect. Dis. 0:0-0(2000).  
DR EMBL: AF219292; AAF30198.1; -  
DR HSSP: P05888; IAAE.  
DR InterPro: IPR001878; znf\_CCHC.  
DR Pfam: PF00098; zif\_CCHC; 1.  
DR SMART: SM00343; znf\_C2HC; 1.  
FT NON\_TER 1 1  
FT NON\_TER 54 54  
SQ SEQUENCE 54 AA; 6000 MW; 9AC21F799195F4AC CRC64;

Query Match 79.5%; Score 35; DB 15; Length 54;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KAKVLAEEA 9  
Db 5 KAKVLAEEA 12

## RESULT 11

ID 09J140 PRELIMINARY: PRT: 55 AA.  
AC 09J140;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Gag protein (Fragment).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FIN8907;  
RA Lillsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
RA Leinikki P., Salminen M.O.;  
RT \*Analysis of HIV-1 genetic subtypes in Finland reveals good  
RT correlation between molecular and epidemiological data.\*;  
RL Scand. J. Infect. Dis. 0:0-0(2000).  
DR EMBL: AF219333; AAF30239.1; -  
DR HSSP: P05888; IAAE.  
DR InterPro: IPR001878; znf\_CCHC.  
DR Pfam: PF00098; zif\_CCHC; 1.  
DR SMART: SM00343; znf\_C2HC; 1.  
FT NON\_TER 1 1  
FT NON\_TER 55 55  
SQ SEQUENCE 55 AA; 6077 MW; 2729CB40CC66804A CRC64;

Query Match 79.5%; Score 35; DB 15; Length 55;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KAKVLAEEA 9  
Db 6 KAKVLAEEA 13

## RESULT 12

ID 09J117 PRELIMINARY: PRT: 56 AA.  
AC 09J117;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Gag protein (Fragment).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FIN9210;  
RA Lillsola K., Holmstrom P., Laukkanen T., Brummer-Korvenkontio H.,  
RA Leinikki P., Salminen M.O.;  
RT \*Analysis of HIV-1 genetic subtypes in Finland reveals good  
RT correlation between molecular and epidemiological data.\*;  
RL Scand. J. Infect. Dis. 0:0-0(2000).  
DR EMBL: AF219356; AAF30262.1; -  
DR HSSP: P05888; IAAE.  
DR InterPro: IPR001878; znf\_CCHC.  
DR Pfam: PF00098; zif\_CCHC; 1.  
DR SMART: SM00343; znf\_C2HC; 1.  
FT NON\_TER 1 1  
FT NON\_TER 56 56  
SQ SEQUENCE 56 AA; 6191 MW; 0DA57152AD3E4482 CRC64;

Query Match 79.5%; Score 35; DB 15; Length 56;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
 |||||  
 DB 6 KARVLAEA 13

## RESULT 13

09J144 PRELIMINARY; PRT; 58 AA.  
 AC 09J144:  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 GN Gag protein (Fragment).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FIN9206;  
 RA Litsola K., Holmstrom P., Laukkanen T., Brunner-Korvenkontio H.,  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data.";  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL; AF219329; AAF30235.1; -  
 DR HSSP; P05888; IAAF  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; zf\_CCHC; 1.  
 DR SMART; SM00343; Znf\_C2HC; 1.  
 FT NON\_TER 1 1  
 FT SEQUENCE 58 AA; 6414 MW; 86845037A48AE958 CRC64;

Query Match 79.5%; Score 35; DB 15; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
 |||||  
 DB 6 KARVLAEA 13

## RESULT 14

09J138 PRELIMINARY; PRT; 58 AA.  
 AC 09J138:  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 GN Gag protein (Fragment).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FIN93173;  
 RA Litsola K., Holmstrom P., Laukkanen T., Brunner-Korvenkontio H.,  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data.";  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL; AF219335; AAF30241.1; -  
 DR HSSP; P05888; IAAF  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; zf\_CCHC; 1.  
 DR SMART; SM00343; Znf\_C2HC; 1.  
 FT NON\_TER 1 1  
 FT SEQUENCE 58 AA; 6313 MW; F3A2592D06632B69 CRC64;

Query Match 79.5%; Score 35; DB 15; Length 58;

Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
 |||||  
 DB 6 KARVLAEA 13

## RESULT 15

09J198 PRELIMINARY; PRT; 61 AA.  
 AC 09J198:  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 GN Gag protein (Fragment).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FIN93214;  
 RA Litsola K., Holmstrom P., Laukkanen T., Brunner-Korvenkontio H.,  
 RT "Analysis of HIV-1 genetic subtypes in Finland reveals good  
 correlation between molecular and epidemiological data.";  
 RL Scand. J. Infect. Dis. 0:0-0(2000).  
 DR EMBL; AF219275; AAF30181.1; -  
 DR HSSP; P05888; IAAF  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00098; zf\_CCHC; 1.  
 DR SMART; SM00343; Znf\_C2HC; 1.  
 FT NON\_TER 1 1  
 FT SEQUENCE 61 AA; 6951 MW; D2A2F472E0ECP696 CRC64;

Query Match 79.5%; Score 35; DB 15; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
 |||||  
 DB 1 KARVLAEA 8

Search completed: March 13, 2003, 18:51:57  
 Job time : 21.7273 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:40:28 ; Search time 5.27273 Seconds  
(without alignments)  
78.662 Million cell updates/sec

Title: US-10-050-200-5  
Perfect score: 44  
Sequence: 1 EKARVLAEMA 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	86.4	557	1 YP85_MYCTU	Q50636 mycobacteri
2	36	81.8	298	1 ARGB_NEIMA	Q91uk2 neisseria m
3	36	81.8	298	1 ARGB_NEIMB	Q91uk2 neisseria m
4	35	79.5	388	1 GAG_HVIW2	P05889 human immun
5	35	79.5	492	1 GAG_HVIU4	P24736 human immun
6	35	79.5	496	1 GAG_HVIU4	P18800 human immun
7	35	79.5	498	1 GAG_HVIU4	P20889 human immun
8	35	79.5	499	1 GAG_HVIC4	P05887 human immun
9	35	79.5	499	1 GAG_HVIC4	P04592 human immun
10	35	79.5	499	1 GAG_HVIC4	P04591 human immun
11	35	79.5	499	1 GAG_HVIC4	P12494 human immun
12	35	79.5	499	1 GAG_HVIC4	Q70692 human immun
13	35	79.5	499	1 GAG_HVIC4	P12493 human immun
14	35	79.5	499	1 GAG_HVIC4	P35962 human immun
15	35	79.5	500	1 GAG_HVIC4	P12495 human immun
16	35	79.5	501	1 GAG_HVIC4	P03349 human immun
17	35	79.5	503	1 GAG_HVIC4	P20873 human immun
18	35	79.5	504	1 GAG_HVIC4	P04594 human immun
19	35	79.5	506	1 GAG_HVIC4	P05888 human immun
20	35	79.5	508	1 GAG_HVIC4	P17282 chimpanzee
21	35	79.5	511	1 GAG_HVIC4	P03347 human immun
22	35	79.5	511	1 GAG_HVIC4	P04593 human immun
23	35	79.5	511	1 GAG_HVIC4	P03348 human immun
24	35	79.5	511	1 GAG_HVIC4	P03349 human immun
25	35	79.5	511	1 GAG_HVIC4	P10182 pseudomonas
26	35	79.5	500	1 GAG_HVIC4	P05890 human immun
27	35	79.5	785	1 MUS2_BACHD	Q9K880 bacillus ha
28	35	79.5	808	1 SEAZ_MYCTU	Q50612 mycobacteri
29	35	79.5	1163	1 RP0D_PEA	P12227 plasmidativ
30	35	79.5	1224	1 RP0D_PEA	P41606 plasmidativ
31	35	79.5	1361	1 RP0D_PEA	P11704 plasmidativ
32	35	79.5	1376	1 RP0D_PEA	P56764 arabinidopsis
33	35	79.5	1384	1 RP0D_PEA	Q91uk5 sinapis alb

34	34	77.3	1388	1 RP0D_TOBAC	P38550 nicotiana t
35	34	77.3	1479	1 RP0D_WHEAT	Q9K993 triticum ae
36	34	77.3	1513	1 RP0D_ORESA	P12909 oryza sativ
37	34	77.3	1527	1 RP0D_MAIZE	P16025 zea mays (m
38	33	75.0	236	1 UL51_PVKKA	Q85227 pseudorabie
39	33	75.0	283	1 AROE_METHA	Q26344 methanobact
40	33	75.0	386	1 RPA2_METVA	P14247 methanococc
41	33	75.0	393	1 RPA2_THRECE	P29158 thermococcu
42	33	75.0	397	1 RPA2_PYRAB	Q9V113 pyrococcus
43	33	75.0	397	1 RPA2_PYRHO	Q93777 pyrococcus
44	33	75.0	516	1 RPA2_CHLRE	P12759 chlamydomon
45	33	75.0	1729	1 RRP5_YEAST	Q05022 saccharomyc

## ALIGNMENTS

```

RESULT 1
ID      YP85_MYCTU      STANDARD:      PRT:      557 AA.
AC      Q50636;
DT      01-NOV-1997 (rel. 35, Created)
DT      01-NOV-1997 (rel. 35, Last sequence update)
DT      16-OCT-2001 (rel. 40, Last annotation update)
DE      Hypothetical lipoprotein Rv2585c precursor.
GN      Rv2585c OR MT2662 OR MTCY227.16.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteria; Actinobacteridae;
OX      Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_Taxid=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-H37Rv;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA      Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA      Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA      Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence."
RL      Nature 393:537-544(1998).
[2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-CDC 1551 / Oshkosh;
RA      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA      Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA      Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA      Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA      Bishai W.;
RT      "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT      laboratory strains."
RT      Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
-!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
-!- (Probable).
-!- SIMILARITY: STRONG, TO M.LEPRAE ML0489 AND DCIAC.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb.slb.ch).
-----
EMBL: Z77724; CAB01258.1; -.
DR      EMBL; AE007100; AAK46975.1; -.
DR      TIGR; MT2662; -.
DR      TubercuList; Rv2585c; -.
DR      InterPro; IPR000914; SBP_dac_5.

```

DR Pfam: PF00496; SBP\_bac-5; 2.  
 KW Hypothetical protein; Membrane; Lipoprotein; signal;  
 FT SIGNAL 1 30 POTENTIAL.  
 FT CHAIN 1 557 HYPOTHETICAL LIPOPROTEIN RV2585C.  
 FT LIPID 31 31 N-ACTYL DIGLYCERIDE (PROBABLE).  
 FT CONFLICT 462 462 C->S (IN REF. 2).  
 SO SEQUENCE 557 AA; 54249A870582F742 CRC64;

Query Match 86.4%; Score 38; DB 1; Length 557;  
 Best Local Similarity 80.0%; Pred. No. 4.6;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEA 10  
 DB 500 ERARILAEMA 509

## RESULT 2

ARGB\_NEIMA STANDARD; PRT; 298 AA.  
 AC 09JUK2:  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Acetylglutamate kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-phosphotransferase).  
 GN ARGB OR NMA1275.  
 OS Neisseria meningitidis (serogroup A).  
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=65699;  
 RN 11

RP SEQUENCE FROM N.A.  
 RC STRAIN=22491 / Serogroup A / Serotype 4A;  
 RX MEDLINE=20222556; PubMed=10761919;  
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C., Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T., Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S., Jaisels R., Leather S., Moule S., Mungall K., Quail M.A., Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J., Whitehead S., Spratt B.G., Barrett B.G.;  
 RA "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491";  
 RT Nature 404:502-506(2000).  
 RL

CC -1 CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate -> ADP + N-acetyl-L-glutamate 5-phosphate.  
 CC -1 PATHWAY: Arginine biosynthesis; second step.  
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1 SIMILARITY: BELONGS TO THE ACETYLGLUTAMATE KINASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: AL162755; CAB84528.1; -  
 DR InterPro: IPR001048; Aa\_kinase.  
 DR InterPro: IPR004662; Acglutkinase.  
 DR InterPro: IPR001057; Glu\_5kinase.  
 DR Pfam: PF00696; aakkinase; 1.  
 DR PRINTS: PR00474; GLU5KINASE.  
 DR TIGRFAMs: TIGR00761; argB; 1.  
 KW Arginine biosynthesis; Transferase; Kinase; Complete proteome.  
 SO SEQUENCE 298 AA; 31269 MW; 7A027E52D5BA1A74 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 298;  
 Best Local Similarity 77.8%; Pred. No. 6.3;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEA 9

DB 11 DKARILAEA 19

## RESULT 3

ARGB\_NEIMA STANDARD; PRT; 298 AA.  
 AC 09J2F7:  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Acetylglutamate kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-phosphotransferase).  
 GN ARGB OR NMB1074.  
 OS Neisseria meningitidis (serogroup B).  
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=491;  
 RN 11

RP SEQUENCE FROM N.A.  
 RC STRAIN=MC58 / Serogroup B;  
 RX MEDLINE=20175755; PubMed=10710307;  
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B., Cotton M.D., Utterback T.R., Khouli H., Qin H., Yamathavan J., Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L., Smith H.O., Fraser C.M., Hoxon E.R., Rappuoli R., Venter J.C.;  
 RA "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58";  
 RT Science 287:1809-1815(2000).  
 RL

CC -1 CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate -> ADP + N-acetyl-L-glutamate 5-phosphate.  
 CC -1 PATHWAY: Arginine biosynthesis; second step.  
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1 SIMILARITY: BELONGS TO THE ACETYLGLUTAMATE KINASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: AE002458; AAF41469.1; -  
 DR TIGR: NMB1074; -  
 DR InterPro: IPR001048; Aa\_kinase.  
 DR InterPro: IPR004662; Acglutkinase.  
 DR InterPro: IPR001057; Glu\_5kinase.  
 DR Pfam: PF00696; aakkinase; 1.  
 DR PRINTS: PR00474; GLU5KINASE.  
 DR TIGRFAMs: TIGR00761; argB; 1.  
 KW Arginine biosynthesis; Transferase; Kinase; Complete proteome.  
 SO SEQUENCE 298 AA; 31263 MW; 7A19DEB5B5A1A74 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 298;  
 Best Local Similarity 77.8%; Pred. No. 6.3;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEA 9  
 DB 11 DKARILAEA 19

## RESULT 4

GAG\_HVIW2 STANDARD; PRT; 388 AA.  
 AC P05889:  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)



DE GAG polyprotein [Contains: Core proteins p17, p24, p2, p7, p1, p6]  
 DE (fragment).  
 DE GAG.  
 OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11705;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86235450; PubMed=3012778;  
 RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,  
 RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;  
 RT "Genetic variation in HIV-III/LAV over time in patients with AIDS or  
 RT at risk for AIDS."  
 RL Science 232:1548-1553(1986).  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM  
 CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO  
 CC WAS PERINATALLY INFECTED BY HER MOTHER.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: K03457; AAB12988.1; -.  
 DR HSSP: P03351; 1E1A.  
 DR HIV: K03457; GAGSWMJ2.  
 DR InterPro: IPR000721; Gag\_p24.  
 DR InterPro: IPR000071; Retroviral\_p17.  
 DR Pfam: PF00540; Gag\_p17; 1.  
 DR Pfam: PF00607; Gag\_p24; 1.  
 KM AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;  
 KM Zinc-finger.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).  
 FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).  
 FT CHAIN 363 376 CORE PROTEIN P2.  
 FT CHAIN 377 >388 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 FT NON\_TER 388 388  
 SO SEQUENCE 388 AA; 43314 MW; EF865BE10ECF7804 CRC64;  
 Query Match 79.5%; Score 35; DB 1; Length 388;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KARVLAEA 9  
 Db 358 KARVLAEA 365  
 RESULT 5  
 GAG\_HV1U4  
 ID GAG\_HV1U4 STANDARD; PRT; 492 AA.  
 AC P24736;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polyprotein [Contains: Core proteins p17, p24, p2, p7, p1, p6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)  
 OS (HIV-1).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11703;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91090981; PubMed=2265025;  
 RA Oram J.D., Downing R.G., Roff M., Clegg J.C.S., Servadei D.,  
 RA Carswell J.W.;  
 RT "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic  
 RT diversity from other HIV-1 isolates."  
 RL AIDS Res. Hum. Retroviruses 6:1073-1078(1990).  
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
 CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
 CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
 CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
 CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
 CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
 CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M62320; AAA75018.1; -.  
 DR HSSP: P05888; 1AAF.  
 DR InterPro: IPR000721; Gag\_p24.  
 DR InterPro: IPR000071; Retroviral\_p17.  
 DR InterPro: IPR001878; Znf.CCHC.  
 DR Pfam: PF00098; Zf-CCHC; 2.  
 DR Pfam: PF00540; Gag\_p17; 1.  
 DR Pfam: PF00607; Gag\_p24; 1.  
 DR PRINTS: PR00939; C2HCZNFINGER.  
 DR PRINTS: PR00234; HIVMATRIX.  
 DR SMART: SM00343; Znf.C2HC; 2.  
 DR PROSITE: PS0158; Zf-CCHC; 2.  
 KM AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;  
 KM Zinc-finger; Repeat.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT CHAIN 1 127 CORE PROTEIN P17 (MATRIX PROTEIN).  
 FT CHAIN 128 358 CORE PROTEIN P24 (CORE ANTIGEN).  
 FT CHAIN 359 371 CORE PROTEIN P2.  
 FT CHAIN 372 425 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
 FT CHAIN 426 441 CORE PROTEIN P1.  
 FT CHAIN 442 492 CORE PROTEIN P6.  
 FT ZN\_FING 383 400 CCHC-TYPE 1.  
 FT ZN\_FING 404 421 CCHC-TYPE 2.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 SO SEQUENCE 492 AA; 54791 MW; E4A3B4D63BB6972 CRC64;  
 Query Match 79.5%; Score 35; DB 1; Length 492;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KARVLAEA 9  
 Db 354 KARVLAEA 361  
 RESULT 6  
 GAG\_HV1ND  
 ID GAG\_HV1ND STANDARD; PRT; 496 AA.  
 AC P18800;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE GAG polyprotein [Contains: Core proteins p17, p24, p2, p7, p1, p6].  
 GN GAG.  
 OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).  
 OS Viruses; Retrovirdae; Retroviridae; Lentivirus.

```

OX NCBI_TaxId=11695;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90034200; PubMed=2806917;
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
RT "Nucleotide sequence of HIV-1 NDK: a highly cytopathic strain of the
RT human immunodeficiency virus."
RL Gene 81:275-284(1989).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEIN FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M27323; AAA44668.1; -.
CC FIRM: J00065; FOLJND.
CC DR HSP: P05888; 1A4F.
CC DR HIV: M27323; GAGSNDK.
CC DR InterPro: IPR000721; Gag_p24.
CC DR InterPro: IPR000071; Retroviral_p17.
CC DR InterPro: IPR001878; Znf_CCHC.
CC DR Pfam: PF00098; Zf_CCHC; 2.
CC DR Pfam: PF00540; Gag_p17; 1.
CC DR Pfam: PF00607; Gag_p24; 1.
CC DR PRINTS: PR00399; C2HCZNFINGER.
CC DR PRINTS: PR00234; HIVMATRIX.
CC DR SMART: SM00343; Znf_C2HC; 2.
CC DR PROSITE: PS50158; Zf_CCHC; 2.
CC KW AIDS: Core protein; Polyprotein; Myristate; Phosphorylation;
CC zinc-finger; Repeat.
CC KM zinc-finger; Repeat.
CC FT INIT_MET 0 0 BY SIMILARITY.
CC FT CHAIN 1 128 CORE PROTEIN P17 (MATRIX PROTEIN).
CC FT CHAIN 129 359 CORE PROTEIN P24 (CORE ANTIGEN).
CC FT CHAIN 360 374 CORE PROTEIN P2.
CC FT CHAIN 375 429 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
CC FT CHAIN 430 445 CORE PROTEIN P1.
CC FT CHAIN 446 496 CORE PROTEIN P6.
CC FT CHAIN 387 404 CCHC-TYPE 1.
CC FT ZN_FING 408 425 CCHC-TYPE 2.
CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC SO SEQUENCE 496 AA; 55152 MW; F50A42E42CFC20AA CRC64;

Query Match 79.5%; Score 35; DB 1; Length 496;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 KARVILAE 9
DB 355 KARVILAE 362

```

```

DE GAG polyprotein [contains: Core proteins p17, p24, p2, p7, p1, p6].
GN GAG.
OS Human immunodeficiency virus type 1 (OY1 isolate) (HIV-1).
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11699;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90148544; PubMed=2559749;
RA Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Main-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
RT individual presenting an atypical western blot."
RL Aids 3:707-715(1989).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEIN FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.
CC -1- MISCELLANEOUS: THE OY1 ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC HEALTHY GABONESE INDIVIDUAL.
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M26727; AAA83391.1; -.
CC DR HSP: P05888; 1A4F.
CC DR HIV: M26727; GAGSOY1.
CC DR InterPro: IPR000721; Gag_p24.
CC DR InterPro: IPR000071; Retroviral_p17.
CC DR InterPro: IPR001878; Znf_CCHC.
CC DR Pfam: PF00098; Zf_CCHC; 2.
CC DR Pfam: PF00540; Gag_p17; 1.
CC DR Pfam: PF00607; Gag_p24; 1.
CC DR PRINTS: PR00399; C2HCZNFINGER.
CC DR PRINTS: PR00234; HIVMATRIX.
CC DR SMART: SM00343; Znf_C2HC; 2.
CC DR PROSITE: PS50158; Zf_CCHC; 2.
CC KW AIDS: Core protein; Polyprotein; Myristate; Phosphorylation;
CC zinc-finger; Repeat.
CC KM zinc-finger; Repeat.
CC FT INIT_MET 0 0 BY SIMILARITY.
CC FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
CC FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
CC FT CHAIN 363 376 CORE PROTEIN P2.
CC FT CHAIN 377 430 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
CC FT CHAIN 431 446 CORE PROTEIN P1.
CC FT CHAIN 447 498 CORE PROTEIN P6.
CC FT CHAIN 388 405 CCHC-TYPE 1.
CC FT ZN_FING 409 426 CCHC-TYPE 2.
CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC SO SEQUENCE 498 AA; 55451 MW; BDFE20658BEB20B1 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 498;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 KARVILAE 9
DB 358 KARVILAE 365

```

```

RESULT 7
GAG_HV10Y STANDARD: PRT; 498 AA.
AC P20889;
DT 01-FEB-1991 (Rel. 17; Created)
DT 01-FEB-1994 (Rel. 28; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)

```

```

RESULT 8
GAG_HV1C4 STANDARD: PRT; 499 AA.
AC P05887;
DT 01-NOV-1988 (Rel. 09; Created)

```

DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Gag polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].  
GN Gag.  
OS Human immunodeficiency virus type 1 (CDC-451 isolate) (HIV-1).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11687;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=67041461; Pubmed=3490666;  
RA Desai S.M., Kalyanaraman V.S., Casey J.M., Srinivasan A.,  
RT Andersen P.R., Devare S.G.;  
RT "Molecular cloning and primary nucleotide sequence analysis of a  
RT distinct human immunodeficiency virus isolate reveal significant  
RT divergence in its genomic sequences."  
RT Proc. Natl. Acad. Sci. U.S.A. 83:8380-8384(1986).  
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
CC -1- P17: THE P24 PROTEIN IS PHOSPHORYLATED.  
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M13136; AAA44306.1; -  
DR PIR: A25523; FOVWH4.  
DR HSSP: P05888; 1A.F.  
DR HTV; M13136; GAGSCD45.  
DR InterPro: IPR000721; Gag\_P24.  
DR InterPro: IPR000071; Retrovir\_P17.  
DR InterPro: IPR001878; Znf\_CCHC.  
DR Pfam: PF00098; zf\_CCHC; 2.  
DR Pfam: PF00540; Gag\_P17; 1.  
DR Pfam: PF00607; Gag\_P24; 1.  
DR PRINTS; PR00939; C2HCZNFINGER.  
DR PRINTS; PR00234; HIVMATRIX.  
DR SMART; SM00343; ZNF\_C2HC; 2.  
DR PROSITE; PSS0158; ZF\_CCHC; 2.  
KW AIDS: Core protein; Polypotein; Myristate; Phosphorylation;  
KW Zinc-finger; Repeat.  
FT INIT\_MET 0  
FT CHAIN 1 131  
FT CHAIN 132 362  
FT CHAIN 363 376  
FT CHAIN 377 431  
FT CHAIN 432 447  
FT CHAIN 448 499  
FT ZN\_FING 389 406  
FT ZN\_FING 410 427  
FT LIPID 1 1  
SO SEQUENCE 499 AA; 55796 MW; 023CA76C9C6F22AD CRC64;  
Query Match 79.5%; Score 35; DB 1; Length 499;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 KARVLAEA 9  
Db 358 KARVLAEA 365  
RESULT 9  
GAG\_HV1EL

ID GAG\_HV1EL STANDARD; PRT; 499 AA.  
AC P04592;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Gag polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].  
GN Gag.  
OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=66245056; Pubmed=2424612;  
RA Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;  
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis  
RT of two isolates from African patients."  
RT Cell 46:63-74(1986).  
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
CC -1- P17: THE P24 PROTEIN IS PHOSPHORYLATED.  
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: K03454; AAA44324.1; -  
DR EMBL: A07108; GAA00611.1; -  
DR HSSP; P05888; 1A.F.  
DR HTV; K03454; GAGSELT.  
DR InterPro: IPR000721; Gag\_P24.  
DR InterPro: IPR000071; Retrovir\_P17.  
DR InterPro: IPR001878; Znf\_CCHC.  
DR Pfam: PF00098; zf\_CCHC; 2.  
DR Pfam: PF00540; Gag\_P17; 1.  
DR Pfam: PF00607; Gag\_P24; 1.  
DR PRINTS; PR00939; C2HCZNFINGER.  
DR PRINTS; PR00234; HIVMATRIX.  
DR SMART; SM00343; ZNF\_C2HC; 2.  
DR PROSITE; PSS0158; ZF\_CCHC; 2.  
KW AIDS: Core protein; Polypotein; Myristate; Phosphorylation;  
KW Zinc-finger; Repeat.  
FT INIT\_MET 0  
FT CHAIN 1 131  
FT CHAIN 132 362  
FT CHAIN 363 377  
FT CHAIN 378 431  
FT CHAIN 432 447  
FT CHAIN 448 499  
FT ZN\_FING 390 407  
FT ZN\_FING 411 428  
FT LIPID 1 1  
SO SEQUENCE 499 AA; 55582 MW; 8A1785A39EAD08D CRC64;  
Query Match 79.5%; Score 35; DB 1; Length 499;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 KARVLAEA 9  
Db 358 KARVLAEA 365  
RESULT 10

GAG\_HV1H2  
ID GAG\_HV1H2 STANDARD: PRT: 499 AA.  
AC P04591;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].  
GN GAG.  
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11706;  
RN [1]  
RP MEDLINE=87299196; PubMed=3040055;  
RA Rather L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,  
RA Gallo R.C., Wong-Staal F.;  
RT "Complete nucleotide sequences of functional clones of the AIDS  
RT virus."  
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).  
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL: K03455; AAB50258.1; -  
DR HSSP: P05888; 1AAB.  
DR HIV: K03455; GAGSHXB2.  
DR InterPro: IPR000721; Retroviral\_P17.  
DR InterPro: IPR001878; Znf\_CCHC.  
DR Pfam: PF00098; Zf-CCHC; 2.  
DR Pfam: PF00540; Gag\_P17; 1.  
DR Pfam: PF00607; Gag\_P24; 1.  
DR PRINTS: PR00939; C2HCZNFINGER.  
DR PRINTS: PR00234; HIVMATRIX.  
DR SMART: SM00343; ZNF\_C2HC; 2.  
DR PROSITE: PS50158; ZF\_CCHC; 2.  
DR AIDS: Core protein; Polyprotein; Myristate; Phosphorylation;  
KW Zinc-finger; Repeat.  
FT INIT\_MER 0 0 BY SIMILARITY.  
FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).  
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).  
FT CHAIN 363 376 CORE PROTEIN P2.  
FT CHAIN 377 431 CORE PROTEIN P7 (Nucleocapsid protein).  
FT CHAIN 432 447 CORE PROTEIN P1.  
FT CHAIN 448 489 CORE PROTEIN P6.  
FT ZN\_FING 389 406 CCHC-TYPE 1.  
FT ZN\_FING 410 427 CCHC-TYPE 2.  
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
SQ SEQUENCE 499 AA: 55798 MW: 774C384D6EACB108 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 499;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVILAE 9  
|||||||  
DB 358 KARVILAE 365

RESULT 11  
GAG\_HV1J3  
ID GAG\_HV1J3 STANDARD: PRT: 499 AA.  
AC P12494;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].  
GN GAG.  
OS Human immunodeficiency virus type 1 (JH3 isolate) (HIV-1).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11694;  
RN [1]  
RP MEDLINE=89352108; PubMed=2669897;  
RA Komyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.,  
RT "Nucleotide sequences of gag and env genes of a Japanese isolate of  
RT HIV-1 and their expression in bacteria."  
RL AIDS Res. Hum. Retroviruses 5:411-419(1989).  
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL: M21137; AAB03522.1; -  
DR HSSP: P05888; 1AAB.  
DR HIV: M21137; GAGSHJ3.  
DR InterPro: IPR000721; Gag\_P24.  
DR InterPro: IPR001878; Retroviral\_P17.  
DR Pfam: PF00098; Zf-CCHC; 2.  
DR Pfam: PF00540; Gag\_P17; 1.  
DR Pfam: PF00607; Gag\_P24; 1.  
DR PRINTS: PR00939; C2HCZNFINGER.  
DR PRINTS: PR00234; HIVMATRIX.  
DR SMART: SM00343; ZNF\_C2HC; 2.  
DR PROSITE: PS50158; ZF\_CCHC; 2.  
DR AIDS: Core protein; Polyprotein; Myristate; Phosphorylation;  
KW Zinc-finger; Repeat.  
FT INIT\_MER 0 0 BY SIMILARITY.  
FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).  
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).  
FT CHAIN 363 376 CORE PROTEIN P2.  
FT CHAIN 377 431 CORE PROTEIN P7 (Nucleocapsid protein).  
FT CHAIN 432 447 CORE PROTEIN P1.  
FT CHAIN 448 489 CORE PROTEIN P6.  
FT ZN\_FING 389 406 CCHC-TYPE 1.  
FT ZN\_FING 410 427 CCHC-TYPE 2.  
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
SQ SEQUENCE 499 AA: 55725 MW: 52E281D4A24AE9 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 499;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVILAE 9  
|||||||  
DB 358 KARVILAE 365

RESULT 12  
GAG\_HV11W  
ID GAG\_HV11W STANDARD: PRT: 499 AA.  
AC 070622;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].  
GN GAG.  
OS Human immunodeficiency virus type 1 (HM12.3 isolate) (HIV-1).  
OC Viruses; Retrovirda; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=82834;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=95127297; PubMed=7826699;  
RA Reltz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,  
RA Shaw G.M., Kong L.T., Weiss S.H., Waters D., Gallo R.C., Blattner W.;  
RT "Vital variability and serum antibody response in a laboratory worker  
RT infected with HIV type 1 (HIV type IIB)."  
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
CC -1- P17: THE P24 PROTEIN IS PHOSPHORYLATED.  
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: U12055; AAA76686.1; -  
DR HSP: P05888; 1AAE.  
DR InterPro: IPR000721; Gag\_p24.  
DR InterPro: IPR000071; Retrovir\_p17.  
DR InterPro: IPR001878; Znf\_CCHC.  
DR Pfam: PF00098; zf\_CCHC; 2.  
DR Pfam: PF00540; Gag\_p17; 1.  
DR Pfam: PF00607; Gag\_p24; 1.  
DR PRINTS: PR00939; C2HC2NFRNGER.  
DR PRINTS: PR00234; HIVMATRIX.  
DR SMART: SM00343; Znf\_C2HC; 2.  
DR PROSITE: PS50158; Zf\_CCHC; 2.  
DR AIDS: Core protein: Polyprotein; Myristate; Phosphorylation;  
KW AIDS; Core protein: Polyprotein; Myristate; Phosphorylation;  
KW Zinc-finger; Repeat.  
FT INIT\_MER 0  
FT CHAIN 1 131 BY SIMILARITY.  
FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).  
FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).  
FT CHAIN 377 431 CORE PROTEIN P2 (CORE ANTIGEN).  
FT CHAIN 432 447 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
FT CHAIN 448 499 CORE PROTEIN P1.  
FT CHAIN 448 499 CORE PROTEIN P6.  
FT ZN\_FING 389 406 CCHC-TYPE 1.  
FT ZN\_FING 410 427 CCHC-TYPE 2.  
FT LIPID 1 MYRISTATE (BY SIMILARITY).  
SQ SEQUENCE 499 AA: 55772 MW: AAZF683546EDC0A9 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 499;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVILAE 9  
DB 358 KARVILAE 365

RESULT 13  
GAG\_HV1N5  
ID GAG\_HV1N5 STANDARD: PRT: 499 AA.  
AC P12493;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].  
GN GAG.  
OS Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).  
OC Viruses; Retrovirda; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11698;  
RN [1]  
RP SEQUENCE FROM N.A. (CLONE PNL4-3).  
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;  
RA Subltted (JUN-1988) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
CC -1- P17: THE P24 PROTEIN IS PHOSPHORYLATED.  
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: M1921; AAA44987.1; -  
DR HSP: P05888; 1AAE.  
DR HIV: M19921; GAGSNL43.  
DR InterPro: IPR000721; Gag\_p24.  
DR InterPro: IPR000071; Retrovir\_p17.  
DR InterPro: IPR001878; Znf\_CCHC.  
DR Pfam: PF00098; zf\_CCHC; 2.  
DR Pfam: PF00540; Gag\_p17; 1.  
DR Pfam: PF00607; Gag\_p24; 1.  
DR PRINTS: PR00939; C2HC2NFRNGER.  
DR PRINTS: PR00234; HIVMATRIX.  
DR SMART: SM00343; Znf\_C2HC; 2.  
DR PROSITE: PS50158; Zf\_CCHC; 2.  
DR AIDS: Core protein: Polyprotein; Myristate; Phosphorylation;  
KW AIDS; Core protein: Polyprotein; Myristate; Phosphorylation;  
KW Zinc-finger; Repeat.  
FT INIT\_MER 0  
FT CHAIN 1 131 BY SIMILARITY.  
FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).  
FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).  
FT CHAIN 377 431 CORE PROTEIN P2 (CORE ANTIGEN).  
FT CHAIN 432 447 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
FT CHAIN 448 499 CORE PROTEIN P1.  
FT CHAIN 448 499 CORE PROTEIN P6.  
FT ZN\_FING 389 406 CCHC-TYPE 1.  
FT ZN\_FING 410 427 CCHC-TYPE 2.  
FT LIPID 1 MYRISTATE (BY SIMILARITY).  
SQ SEQUENCE 499 AA: 55687 MW: CBEC1102FE2C1E2 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 499;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVILAE 9  
DB 358 KARVILAE 365

OY 2 KARVILAE 9  
DB 358 KARVILAE 365

RESULT 14  
GAG\_HV1Y2  
ID GAG\_HV1Y2 STANDARD: PRT: 499 AA.

```

AC P35962;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GAG polypeptide [Contains: Core proteins p17, p24, p2, p7, p1, p6].
GN GAG.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Burgess C.D., Price R.W., Sharp P.M., Hahn B.H.,
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M93258; -; NOT_ANNOTATED_CDS.
DR PIR: A44001; A44001.
DR HSSP: P05888; 1AAE.
DR InterPro: IPR000721; Gag_p24.
DR InterPro: IPR000071; Retrovir_p17.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00098; Zf_CCHC; 2.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS: PR00939; C2HCZNFINGER.
DR SMART: SM00343; ZNF_C2HC; 2.
DR PROSITE: PS50158; ZF_CCHC; 2.
KW AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
KW Zinc-finger; Repeat.
FT INT_MET 0
FT CHAIN 1 131 BY SIMILARITY.
FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 377 431 CORE PROTEIN P2 (NUCLEOCAPSID PROTEIN).
FT CHAIN 432 447 CORE PROTEIN P1.
FT CHAIN 448 499 CORE PROTEIN P6.
FT ZN_FING 389 406 CCHC-TYPE 1.
FT ZN_FING 410 427 CCHC-TYPE 2.
FT LIPID 1 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 499 AA; 55660 MW; 278E65F5405CD99 CRC64;
Query Match 79.5%; Score 35; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 KARVLAEA 9
DB 358 KARVLAEA 365

```

RESULT 15

```

GAG_HV122
ID GAG_HV122 STANDARD; PRT; 500 AA.
AC P12495;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GAG polypeptide [Contains: Core proteins p17, p24, p2, p7, p1, p6].
GN GAG.
OS Human immunodeficiency virus type 1 (Z2/CDC-234 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11683;
RN [1]
RP SEQUENCE FROM N.A.
RX Theodore T., Buckler-White A.;
RA Submitted (NOV-1989) to the HIV data bank.
RT -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
RT ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
RT REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
RT MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
RT RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
RT GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
RT BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M22639; AAA45365.1; -
DR HSSP: P05888; 1AAE.
DR HIV: M22639; GAGS2226.
DR InterPro: IPR000721; Gag_p24.
DR InterPro: IPR000071; Retrovir_p17.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00098; Zf_CCHC; 2.
DR Pfam: PF00540; Gag_p17; 1.
DR Pfam: PF00607; Gag_p24; 1.
DR PRINTS: PR00939; C2HCZNFINGER.
DR SMART: SM00343; ZNF_C2HC; 2.
DR PROSITE: PS50158; ZF_CCHC; 2.
KW AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
KW Zinc-finger; Repeat.
FT INT_MET 0
FT CHAIN 1 132 BY SIMILARITY.
FT CHAIN 133 363 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 364 378 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 379 433 CORE PROTEIN P2 (NUCLEOCAPSID PROTEIN).
FT CHAIN 434 449 CORE PROTEIN P1.
FT CHAIN 450 500 CORE PROTEIN P6.
FT ZN_FING 391 408 CCHC-TYPE 1.
FT ZN_FING 412 429 CCHC-TYPE 2.
FT LIPID 1 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 500 AA; 55546 MW; B27C2609A92FEB2A CRC64;
Query Match 79.5%; Score 35; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 KARVLAEA 9
DB 359 KARVLAEA 366

```

Search completed: March 13, 2003, 18:49:53  
 Job time : 5.27273 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:49:22 ; Search time 9.4 Seconds

(without alignments)  
53.938 Million cell updates/sec

Title: US-10-050-200-4

Perfect score: 51

Sequence: 1 ADLSEFKSOEL 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppa/PCUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	68.6	361	9	US-09-925-299-804
2	35	68.6	361	10	US-09-925-299-804
3	34	66.7	402	9	US-09-987-021-4
4	34	66.7	402	10	US-09-957-485-4
5	33	64.7	378	10	US-09-816-664-7
6	33	64.7	659	9	US-10-101-464A-75
7	33	64.7	1531	10	US-09-876-889-347
8	33	64.7	1531	10	US-09-998-598-253
9	32	62.7	314	10	US-09-318-271-2
10	31	60.8	206	10	US-09-731-872-352
11	31	60.8	216	9	US-09-984-245-172
12	31	60.8	229	9	US-09-992-598-373
13	31	60.8	229	9	US-09-989-293A-373
14	31	60.8	229	9	US-09-989-735-373
15	31	60.8	229	9	US-09-990-444-373
16	31	60.8	229	9	US-09-989-730-373
17	31	60.8	229	9	US-09-990-436-373
18	31	60.8	229	9	US-09-991-181-373
19	31	60.8	229	9	US-09-993-687-373

20	31	60.8	229	9	US-09-989-734-373	Sequence 373, App
21	31	60.8	229	9	US-10-028-072-472	Sequence 472, App
22	31	60.8	229	9	US-09-997-653-373	Sequence 373, App
23	31	60.8	229	9	US-09-993-667-373	Sequence 373, App
24	31	60.8	229	9	US-10-121-049-472	Sequence 472, App
25	31	60.8	229	9	US-10-123-904-472	Sequence 472, App
26	31	60.8	229	9	US-10-140-470-472	Sequence 472, App
27	31	60.8	229	9	US-09-990-438-373	Sequence 373, App
28	31	60.8	229	9	US-09-990-562-373	Sequence 373, App
29	31	60.8	229	9	US-09-997-428-373	Sequence 373, App
30	31	60.8	229	9	US-09-997-666-373	Sequence 373, App
31	31	60.8	229	9	US-10-175-746-472	Sequence 472, App
32	31	60.8	229	9	US-10-176-918-472	Sequence 472, App
33	31	60.8	229	9	US-10-176-921-472	Sequence 472, App
34	31	60.8	229	9	US-09-990-711-373	Sequence 373, App
35	31	60.8	229	9	US-10-137-865-472	Sequence 472, App
36	31	60.8	229	9	US-10-142-431-472	Sequence 472, App
37	31	60.8	229	9	US-10-142-474-472	Sequence 472, App
38	31	60.8	229	9	US-10-143-114-472	Sequence 472, App
39	31	60.8	229	9	US-09-989-726-373	Sequence 373, App
40	31	60.8	229	9	US-10-140-002-472	Sequence 472, App
41	31	60.8	229	9	US-09-990-437-373	Sequence 373, App
42	31	60.8	229	9	US-09-998-156-373	Sequence 373, App
43	31	60.8	229	9	US-10-142-419-472	Sequence 472, App
44	31	60.8	229	10	US-09-989-722-373	Sequence 373, App
45	31	60.8	229	10	US-09-989-723-373	Sequence 373, App

#### ALIGNMENTS

RESULT 1  
US-09-925-299-804  
; Sequence 804, Application US/09925299  
; Publication No. US20030040617A9  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05683  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 804  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-299-804

Query Match 68.6% Score 35; DB 9; Length 361;  
Best Local Similarity 70.0% Pred. No. 25;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DLSEFKSOEL 11  
| | | | | | | |  
Db 317 DRSEFKSOEL 326

RESULT 2  
US-09-925-299-804  
; Sequence 804, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05683

;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 60/124,270  
;; PRIOR FILING DATE: 1999-03-12  
;; NUMBER OF SEQ ID NOS: 1556  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO: 804  
;; LENGTH: 361  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-925-299-804

Query Match 68.6%; Score 35; DB 10; Length 361;  
Best Local Similarity 70.0%; Pred. No. 25;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DLSSFKSQE 11  
DB 317 DRSSFKQOE 326

## RESULT 3

US-09-987-021-4  
;; Sequence 4, Application US/09987021  
;; Patent No. US20020165147A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Yeees, et al.  
;; TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen Activator  
;; FILE REFERENCE: PF336P2  
;; CURRENT APPLICATION NUMBER: US/09/987,021  
;; PRIOR FILING DATE: 2001-11-13  
;; PRIOR APPLICATION NUMBER: 09/957,485  
;; PRIOR FILING DATE: 2001-09-21  
;; PRIOR APPLICATION NUMBER: 09/722,292  
;; PRIOR FILING DATE: 2000-11-28  
;; PRIOR APPLICATION NUMBER: 60/247,971  
;; PRIOR FILING DATE: 2000-11-14  
;; PRIOR APPLICATION NUMBER: 09/521,664  
;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 09/348,817  
;; PRIOR FILING DATE: 1999-07-08  
;; PRIOR APPLICATION NUMBER: 60/123,704  
;; PRIOR FILING DATE: 1999-03-10  
;; PRIOR APPLICATION NUMBER: 08/948,997  
;; PRIOR FILING DATE: 1997-10-10  
;; PRIOR APPLICATION NUMBER: 60/028,117  
;; PRIOR FILING DATE: 1996-10-11  
;; NUMBER OF SEQ ID NOS: 18  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO: 4  
;; LENGTH: 402  
;; TYPE: PRT  
;; ORGANISM: Bos taurus  
US-09-987-021-4

Query Match 66.7%; Score 34; DB 9; Length 402;  
Best Local Similarity 70.0%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ADLSSFKSQE 10  
DB 327 ADPSSFSQOE 336

## RESULT 4

US-09-957-485-4  
;; Sequence 4, Application US/09957485  
;; Patent No. US20020143165A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Human Genome Sciences, Inc. et al.  
;; TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen  
;; TITLE OF INVENTION: Activator  
;; FILE REFERENCE: PF336P1  
;; CURRENT APPLICATION NUMBER: US/09/957,485

;; CURRENT FILING DATE: 2001-09-21  
;; PRIOR APPLICATION NUMBER: US 09/521,664  
;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: US 60/123,704  
;; PRIOR FILING DATE: 1999-03-10  
;; NUMBER OF SEQ ID NOS: 21  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO: 4  
;; LENGTH: 402  
;; TYPE: PRT  
;; ORGANISM: Bos taurus  
US-09-957-485-4

Query Match 66.7%; Score 34; DB 10; Length 402;  
Best Local Similarity 70.0%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ADLSSFKSQE 10  
DB 327 ADPSSFSQOE 336

## RESULT 5

US-09-816-664-7  
;; Sequence 7, Application US/09816664  
;; Patent No. US20020038014A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Meyers, Rachel A.  
;; TITLE OF INVENTION: 26443 and 46873, NOVEL HUMAN  
;; TITLE OF INVENTION: ASPARAGINASE FAMILY MEMBERS AND USES THEREFOR  
;; FILE REFERENCE: 10448-029001  
;; CURRENT APPLICATION NUMBER: US/09/816,664  
;; CURRENT FILING DATE: 2001-03-23  
;; PRIOR APPLICATION NUMBER: US 60/191,973  
;; PRIOR FILING DATE: 2000-03-24  
;; NUMBER OF SEQ ID NOS: 9  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 7  
;; LENGTH: 378  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: consensus sequence  
US-09-816-664-7

Query Match 64.7%; Score 33; DB 10; Length 378;  
Best Local Similarity 66.7%; Pred. No. 65;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 DLSSFKSQE 10  
DB 162 DLSTFETQE 170

## RESULT 6

US-10-101-464A-75  
;; Sequence 75, Application US/10101464A  
;; Publication No. US20030046728A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Stradala, Timothy  
;; APPLICANT: Nieuwenhuizen, Nicolaas  
;; APPLICANT: Higgins, Colleen M.  
;; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
;; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
;; FILE REFERENCE: 11000.1020C2  
;; CURRENT APPLICATION NUMBER: US/10/101,464A  
;; CURRENT FILING DATE: 2002-03-18  
;; PRIOR APPLICATION NUMBER: 09/704,302  
;; PRIOR FILING DATE: 2000-11-01  
;; PRIOR APPLICATION NUMBER: 09/228,986  
;; PRIOR FILING DATE: 1999-01-12  
;; PRIOR APPLICATION NUMBER: 60/162,866  
;; PRIOR FILING DATE: 1999-11-01



; PRIOR APPLICATION NUMBER: PCT/US00/00724  
 ; PRIOR FILING DATE: 2000-01-11  
 ; NUMBER OF SEQ ID NOS: 989  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 75  
 ; LENGTH: 659  
 ; TYPE: PRT  
 ; ORGANISM: Eucalyptus grandis  
 US-10-101-464A-75

Query Match 64.7%; Score 33; DB 9; Length 659;  
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 DLSSFKSOEL 11  
 DB 289 DVATFKHOEL 298

RESULT 7  
 US-09-876-889-347  
 ; Sequence 347, Application US/09876889  
 ; Patent No. US20020076715A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Benson, Darin R.  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: King, Gordon E.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN  
 ; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS  
 ; FILE REFERENCE: 210121.466C3  
 ; CURRENT APPLICATION NUMBER: US/09/876,889  
 ; CURRENT FILING DATE: 2001-06-06  
 ; NUMBER OF SEQ ID NOS: 353  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 347  
 ; LENGTH: 1531  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-876-889-347

Query Match 64.7%; Score 33; DB 10; Length 1531;  
 Best Local Similarity 70.0%; Pred. No. 3e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLSSFKSOEL 11  
 DB 222 DLSSFKSOEL 231

RESULT 8  
 US-09-998-598-2593  
 ; Sequence 2593, Application US/09998598  
 ; Patent No. US20020150922A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stolt, John A.  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Chenault, Ruth A.  
 ; APPLICANT: Meagher, Madelein Joy  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER  
 ; FILE REFERENCE: 210121.561  
 ; CURRENT APPLICATION NUMBER: US/09/998,598  
 ; CURRENT FILING DATE: 2001-11-16  
 ; NUMBER OF SEQ ID NOS: 2606  
 ; SOFTWARE: Corixa Invention Disclosure Database  
 ; SEQ ID NO 2593  
 ; LENGTH: 1531  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-998-598-2593

Query Match 64.7%; Score 33; DB 10; Length 1531;

Best Local Similarity 70.0%; Pred. No. 3e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLSSFKSOEL 11  
 DB 222 DLSSFKSOEL 231

RESULT 9  
 US-09-318-271-2  
 ; Sequence 2, Application US/09318271A  
 ; Patent No. US20020012979A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berry, Alan  
 ; APPLICANT: Running, Jeffrey A.  
 ; APPLICANT: Severson, David K.  
 ; APPLICANT: Burlingame, Richard P.  
 ; TITLE OF INVENTION: "VITAMIN C PRODUCTION IN MICROORGANISMS AND PLANTS"  
 ; FILE REFERENCE: 3161-24  
 ; CURRENT APPLICATION NUMBER: US/09/318,271A  
 ; CURRENT FILING DATE: 1999-05-25  
 ; EARLIER APPLICATION NUMBER: 60/125,073  
 ; EARLIER FILING DATE: 1999-03-17  
 ; EARLIER APPLICATION NUMBER: 60/125,054  
 ; EARLIER FILING DATE: 1999-03-18  
 ; EARLIER APPLICATION NUMBER: 60/088,549  
 ; EARLIER FILING DATE: 1998-06-08  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 314  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-318-271-2

Query Match 62.7%; Score 32; DB 10; Length 314;  
 Best Local Similarity 70.0%; Pred. No. 83;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADLSSFKSOE 10  
 DB 47 ADVESFESOE 56

RESULT 10  
 US-09-731-872-352  
 ; Sequence 352, Application US/09731872  
 ; Patent No. US20020102604A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
 ; APPLICANT: Bougueleret, Lydie  
 ; APPLICANT: Jobert, Severin  
 ; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
 ; FILE REFERENCE: 78 US3 REG  
 ; CURRENT APPLICATION NUMBER: US/09/731,872  
 ; CURRENT FILING DATE: 2000-12-07  
 ; PRIOR APPLICATION NUMBER: US 60/169,629  
 ; PRIOR FILING DATE: 1999-12-08  
 ; PRIOR APPLICATION NUMBER: US 60/187,470  
 ; PRIOR FILING DATE: 2000-03-06  
 ; NUMBER OF SEQ ID NOS: 482  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 352  
 ; LENGTH: 206  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SIGNAL  
 ; LOCATION: -34..-1  
 US-09-731-872-352

Query Match 60.8%; Score 31; DB 10; Length 206;  
 Best Local Similarity 87.5%; Pred. No. 82;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 ADLSSFKS 8  
1111111  
Db 93 ADLSSFKS 100

RESULT 11  
US-09-984-245-172  
; Sequence 172, Application US/09984245  
; Patent No. US20020165374A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 87 Human Secreted Proteins  
; FILE REFERENCE: P2004P1  
; CURRENT APPLICATION NUMBER: US/09/984,245  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 09/134,707  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/05311  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: US 60/041,277  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/042,344  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,276  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,281  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/048,094  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,350  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,188  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,135  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/050,937  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,187  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,099  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,352  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,186  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,069  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,095  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,131  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,096  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,355  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,160  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,351  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,154  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/054,804  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: US 60/056,370  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: US 60/060,862  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 343  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 172  
; LENGTH: 216

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-984-245-172

Query Match 60.8%; Score 31; DB 9; Length 216;  
Best Local Similarity 87.5%; Pred. No. 87;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ADLSSFKS 8  
1111111  
Db 80 ADLSSFKS 87

RESULT 12  
US-09-992-598-373  
; Sequence 373, Application US/09992598  
; Patent No. US20020160384A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gertlsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C20  
; CURRENT APPLICATION NUMBER: US/09/992,598  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02



QY 1 ADLSFRS 8  
Db 93 ADLSFRS 100

RESULT 13  
US-09-989-293A-373  
Sequence 373, Application US/09989293A  
Patent No. US20020177164A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Bolstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoul, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC66  
CURRENT APPLICATION NUMBER: US/09/989,293A  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026

PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22

;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 60.8%; Score 31; DB 9; Length 229;  
Best Local Similarity 87.5%; Pred. No. 93;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSSEFS 8  
DB 93 ADLSSEFS 100

RESULT 14

US-09-989-735-737  
;; Sequence 373, Application US/09989735  
;; Publication No. US20020193299A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ashkenazi, Avi J.  
;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Eaton, Dan L.  
;; APPLICANT: Ferrara, Napoleone  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gertlisen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Kljavin, Ivar J.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James  
;; APPLICANT: Paoni, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Watanabe, Colin K.  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Zhang, Zemin  
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
;; TITLE OF INVENTION: Acids Encoding the Same  
;; FILE REFERENCE: P2730P1C61  
;; CURRENT APPLICATION NUMBER: US/09/989,735  
;; CURRENT FILING DATE: 2001-11-19  
;; PRIOR APPLICATION NUMBER: 60/049787  
;; PRIOR FILING DATE: 1997-06-16  
;; PRIOR APPLICATION NUMBER: 60/062250  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/065186  
;; PRIOR FILING DATE: 1997-11-12  
;; PRIOR APPLICATION NUMBER: 60/065311  
;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/075945  
;; PRIOR FILING DATE: 1998-02-25  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/087607  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087609  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087759  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087827  
;; PRIOR FILING DATE: 1998-06-03  
;; PRIOR APPLICATION NUMBER: 60/088021  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088025  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088028  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088029  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088030  
;; PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088033  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088326  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088167  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088202  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088212  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088217  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088655  
 PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: 60/088734  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088738  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088742  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088824  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088826  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088861  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088876  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089105  
 PRIOR FILING DATE: 1998-06-12  
 PRIOR APPLICATION NUMBER: 60/089440  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089512  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089514  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089538  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089598  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089600  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089653  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089801  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089908  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090246  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090252  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090254  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090355

PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090435  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090444  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090472  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090535  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090540  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090542  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090557  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090676  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090678  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090690  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090694  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090695  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090696  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090862  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091478  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091544  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091626  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091633  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091978  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/092182  
 PRIOR FILING DATE: 1998-07-09

Query Match 60.8%; Score 31; DB 9; Length 229;  
 Best Local Similarity 87.5%; Pred. No. 93;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLSSEFS 8  
 |||||  
 Db 93 ADLSSEFS 100

RESULT 15  
 US-09-990-444-373  
 ; Sequence 373, Application US/09990444  
 ; Publication No. US20020193300A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC19  
CURRENT APPLICATION NUMBER: US/09/990,444  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202

PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 60.8%; Score 31; DB 9; Length 229;  
Best Local Similarity 87.5%; Pred. No. 93;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLSEFS 8  
|||  
Db 93 ADLSEFS 100

Search completed: March 13, 2003, 18:54:46  
Job time : 10.4 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:52 ; Search time 9.8 Seconds  
(without alignments)  
33.026 Million cell updates/sec

Title: US-10-050-200-4

Perfect score: 51

Sequence: 1 ADLSSFKSQEL 11

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5  
262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents RA: \*  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/1aa/PCUTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	166	1	US-08-365-103B-14
2	51	100.0	167	1	US-08-365-103B-12
3	51	100.0	330	1	US-08-365-103B-10
4	51	100.0	321	1	US-08-365-103B-8
5	34	66.7	402	3	US-08-948-997-4
6	34	66.7	402	3	US-09-348-817A-4
7	33	64.7	174	1	US-08-261-825-2
8	33	64.7	174	2	US-08-719-124-2
9	33	64.7	174	5	PCT-US95-07748A-2
10	33	64.7	659	4	US-09-228-986-75
11	32	62.7	591	2	US-08-869-402-1
12	32	62.7	1019	1	US-08-296-014A-4
13	32	62.7	1019	2	US-08-596-405-4
14	32	62.7	1019	2	US-08-877-620-4
15	32	62.7	1083	1	US-08-296-014A-2
16	32	62.7	1083	2	US-08-596-405-2
17	32	62.7	1083	2	US-08-877-620-2
18	31	60.8	106	4	US-08-905-223-413
19	31	60.8	187	4	US-09-535-521-17
20	31	60.8	196	4	US-09-247-155-163
21	31	60.8	208	4	US-09-535-521-20
22	31	60.8	292	4	US-09-535-521-2
23	31	60.8	292	4	US-09-535-521-5
24	31	60.8	3031	1	US-07-689-008-2
25	30	58.8	204	1	US-08-292-945-8
26	30	58.8	204	1	US-08-252-073A-8
27	30	58.8	204	5	PCT-US93-12074-8

28	30	58.8	247	4	US-09-370-838-112	Sequence 112, App
29	30	58.8	630	2	US-08-394-177-5	Sequence 5, Appl
30	30	58.8	630	4	US-08-394-177-5	Sequence 5, Appl
31	30	58.8	630	4	US-09-333-636-5	Sequence 5, Appl
32	30	58.8	1326	4	US-09-147-236-5	Sequence 5, Appl
33	30	58.8	2037	4	US-09-306-998-3	Sequence 3, Appl
34	29	56.9	134	5	PCT-US93-07213-15	Sequence 15, Appl
35	29	56.9	244	1	US-08-910-973-22	Sequence 22, Appl
36	29	56.9	244	4	US-09-499-227-22	Sequence 22, Appl
37	29	56.9	255	4	US-08-800-682-3	Sequence 3, Appl
38	29	56.9	270	4	US-09-134-001C-5016	Sequence 5016, Ap
39	29	56.9	295	4	US-09-134-001C-4059	Sequence 4059, Ap
40	29	56.9	352	4	US-08-800-682-2	Sequence 2, Appl
41	29	56.9	455	3	US-09-221-925-5	Sequence 5, Appl
42	29	56.9	455	3	US-09-221-928-5	Sequence 5, Appl
43	29	56.9	455	4	US-09-221-927-5	Sequence 5, Appl
44	29	56.9	455	4	US-09-221-936-5	Sequence 5, Appl
45	29	56.9	455	4	US-09-221-416-5	Sequence 5, Appl

#### ALIGNMENTS

RESULT 1  
US-08-365-103B-14  
Sequence 14, Application US/08365103B  
Patent No. 5766943  
GENERAL INFORMATION:  
APPLICANT: Lynch, Richard G.  
APPLICANT: Nunez, Raphael D.  
TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23  
TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23  
NUMBER OF INVENTION: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease  
STREET: 801 Grand Ave. Suite 3200  
CITY: Des Moines  
STATE: Iowa  
COUNTRY: United States  
ZIP: 50309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/365,103B  
FILING DATE: 28-DEC-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Nebel, Heidi S.  
REGISTRATION NUMBER: 37,719  
REFERENCE/DOCKET NUMBER: Ulrf N5-24  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (515) 288-3667  
TELEFAX: (515) 288-1338  
INFORMATION FOR SEO ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 166 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-365-103B-14

Query Match 100.0%; Score 51; DB 1; Length 166;

Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADLSSFKSQEL 11  
|||||  
Db 79 ADLSSFKSQEL 89

```
RESULT 2
US-08-365-103B-12
; Sequence 12, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G
; APPLICANT: Nunez, Raphael D.
; TITLE OF INVENTION: DNA Sequences for Soluble Forns of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave., Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Ulrf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-365-103B-12
;
Query Match          100.0%; Score 51; DB 1; Length 167;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADLSFKSOEL 11
      1111111111
DB 80 ADLSFKSOEL 90

RESULT 3
US-08-365-103B-10
; Sequence 10, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G
; APPLICANT: Nunez, Raphael D.
; TITLE OF INVENTION: DNA Sequences for Soluble Forns of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave., Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Ulrf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-365-103B-10
;
Query Match          100.0%; Score 51; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADLSFKSOEL 11
      1111111111
DB 117 ADLSFKSOEL 127

RESULT 4
US-08-365-103B-8
; Sequence 8, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G
; APPLICANT: Nunez, Raphael D.
; TITLE OF INVENTION: DNA Sequences for Soluble Forns of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave., Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Ulrf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-365-103B-8
;
Query Match          100.0%; Score 51; DB 1; Length 321;
```

Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADLSFKSOEL 11  
|||||  
Db 118 ADLSFKSOEL 128

## RESULT 5

US-08-948-997-4  
; Sequence 4, Application US/08948997  
; Patent No. 6008020  
; GENERAL INFORMATION:  
; APPLICANT: HASTINGS, GREGG  
; APPLICANT: COLEMAN, TIM  
; APPLICANT: LAWRENCE, DANIEL  
; TITLE OF INVENTION: BRAIN-ASSOCIATED INHIBITOR OF  
; TITLE OF INVENTION: TISSUE-TYPE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
; STREET: 9410 KEY WEST AVENUE  
; CITY: ROCKVILLE  
; STATE: MD  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/948,997  
; FILING DATE: Oct-10-97  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. ANDERS BROOKES  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PF336  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-948-997-4

Query Match 66.7%; Score 34; DB 3; Length 402;  
Best Local Similarity 70.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ADLSFKSOE 10  
|||||  
Db 327 ADLSFKSOE 336

## RESULT 6

US-09-348-817A-4  
; Sequence 4, Application US/09348817A  
; Patent No. 6191260  
; GENERAL INFORMATION:  
; APPLICANT: Hastings et al.  
; TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen  
; TITLE OF INVENTION: Activator  
; FILE REFERENCE: PF336D1  
; CURRENT APPLICATION NUMBER: US/09/348,817A  
; CURRENT FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: 08/948,997  
; PRIOR FILING DATE: 1997-10-10

PRIOR APPLICATION NUMBER: 60/028,117  
; PRIOR FILING DATE: 1996-10-11  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Bos taurus  
; US-09-348-817A-4

Query Match 66.7%; Score 34; DB 4; Length 402;  
Best Local Similarity 70.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ADLSFKSOE 10  
|||||  
Db 327 ADLSFKSOE 336

## RESULT 7

US-08-261-825-2  
; Sequence 2, Application US/08261825  
; Patent No. 5558993  
; GENERAL INFORMATION:  
; APPLICANT: Champion, Cheryl I.  
; APPLICANT: Lovett, Michael A.  
; APPLICANT: Haake, David A.  
; APPLICANT: Miller, James N.  
; APPLICANT: Blanco, David R.  
; TITLE OF INVENTION: CLONED Borrelia burgdorferi VIRULENCE  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Spensley Horn Judas & Lubitz  
; STREET: 1880 Century Park East, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90067  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/261,825  
; FILING DATE: 17-JUN-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: TURAKIN, LISA A.  
; REGISTRATION NUMBER: P-38,347  
; REFERENCE/DOCKET NUMBER: PD3516  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 455-5100  
; TELEFAX: (619) 455-5110  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 174 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-261-825-2

Query Match 64.7%; Score 33; DB 1; Length 174;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 DLSSFKSOEL 11  
|||||  
Db 17 DLSSFKSOE 26

## RESULT 8

US-08-719-124-2  
; Sequence 2, Application US/08719124  
; Patent No. 5854395  
; GENERAL INFORMATION:  
; APPLICANT: Champion, Cheryl I.  
; APPLICANT: Lovett, Michael A.  
; APPLICANT: Haake, David A.  
; APPLICANT: Miller, James N.  
; APPLICANT: Blanco, David R.  
; TITLE OF INVENTION: CLONED Borrelia burgdorferi VIRULENCE  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Jubas & Lubitz  
; STREET: 1880 Century Park East, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90067  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; APPLICATION NUMBER: US/08/719,124  
; FILING DATE: 24-SEP-1996  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/261,825  
; FILING DATE: 17-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: TUMARKIN, LISA A.  
; REGISTRATION NUMBER: P-38,347  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 455-5100  
; TELEFAX: (619) 455-5110  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 174 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-719-124-2  
  
Query Match 64.7%; Score 33; DB 2; Length 174;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
OY 2 DLSSFKSOEL 11  
DB 17 DLSAFMSQDI 26  
  
RESULT 9  
PCT-US95-07748A-2  
; Sequence 2, Application PC/TUS9507748A  
; GENERAL INFORMATION:  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: CLONED Borrelia burgdorferi VIRULENCE  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; SOFTWARE: IBM PC compatible  
; CURRENT APPLICATION DATA:

OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/07748A  
; FILING DATE: 16-JUN-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HAILE, Ph.D., LISA A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07419/013W01 (PD3516)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 678-5070  
; TELEFAX: (619) 678-5099  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 174 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-07748A-2  
  
Query Match 64.7%; Score 33; DB 5; Length 174;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
OY 2 DLSSFKSOEL 11  
DB 17 DLSAFMSQDI 26  
  
RESULT 10  
US-09-228-986-75  
; Sequence 75, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions isolated from Plant Cells  
; TITLE OF INVENTION: and their use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 75  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-09-228-986-75  
  
Query Match 64.7%; Score 33; DB 4; Length 659;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
OY 2 DLSSFKSOEL 11  
DB 289 DVATFKHQEL 298  
  
RESULT 11  
US-08-889-402-1  
; Sequence 1, Application US/08889402  
; Patent No. 5811288  
; GENERAL INFORMATION:  
; TITLE OF INVENTION: NOVEL ISOFORM GENE FOR FOCAL ADHESION  
; TITLE OF INVENTION: PROTEIN PAXILLIN  
; NUMBER OF SEQUENCES: 7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/889,402  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 178334/1996  
FILING DATE: 08-JUL-1996  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 591 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
CELL TYPE: Monoblastic cell and placenta  
CELL LINE: U937  
US-08-889-402-1

Query Match 62.7%; Score 32; DB 2; Length 591;  
Best Local Similarity 63.6%; Pred. No. 2e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ADLSPKSOEL 11  
DB 271 ASLSPKIQDL 281

RESULT 12  
US-08-296-014A-4  
Sequence 4, Application US/08296014A  
Patent No. 5716834  
GENERAL INFORMATION:  
APPLICANT: Ding, Jeak Ling  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1781-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1019 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-296-014A-4

Query Match 62.7%; Score 32; DB 1; Length 1019;  
Best Local Similarity 54.5%; Pred. No. 3.6e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADLSPKSOEL 11  
DB 404 SDLNGKSEEL 414

RESULT 13  
US-08-596-405-4  
Sequence 4, Application US/08596405  
Patent No. 5856706  
GENERAL INFORMATION:  
APPLICANT: Ding, Jeak Ling  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1781-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1019 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-596-405-4

Query Match 62.7%; Score 32; DB 2; Length 1019;  
Best Local Similarity 54.5%; Pred. No. 3.6e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADLSPKSOEL 11  
DB 404 SDLNGKSEEL 414

RESULT 14  
US-08-877-620-4  
Sequence 4, Application US/08877620  
Patent No. 5985590  
GENERAL INFORMATION:  
APPLICANT: Ding, Jeak Ling  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1781-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1019 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-877-620-4

Query Match 62.7%; Score 32; DB 1; Length 1019;  
Best Local Similarity 54.5%; Pred. No. 3.6e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADLSPKSOEL 11  
DB 404 SDLNGKSEEL 414

RESULT 15  
US-08-877-620-4  
Sequence 4, Application US/08877620  
Patent No. 5985590  
GENERAL INFORMATION:  
APPLICANT: Ding, Jeak Ling  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1781-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1019 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-877-620-4

Query Match 62.7%; Score 32; DB 1; Length 1019;  
Best Local Similarity 54.5%; Pred. No. 3.6e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADLSPKSOEL 11  
DB 404 SDLNGKSEEL 414

STATE: Virginia  
COUNTRY: USA  
ZIP: 22042  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/877,620  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/596,405  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1781-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1019 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-877-620-4

Query Match 62.7%; Score 32; DB 2; Length 1019;  
Best Local Similarity 54.5%; Pred. No. 3.6e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADLSFKSOEL 11  
:||: ||:|  
DB 404 SDLNGIKSEEL 414

RESULT 15  
US-08-296-014A-2  
Sequence 2, Application US/08296014A  
Patent No. 5716834  
GENERAL INFORMATION:  
APPLICANT: Ding, Jeak Ling  
APPLICANT: HO, BOU  
TITLE OF INVENTION: The Cloned Factor C cDNA of the  
TITLE OF INVENTION: Singapore Horseshoe Crab, Carinoscorpius  
NUMBER OF INVENTIONS: 4  
TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: 8110 Gatehouse Road, Suite 500 East  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22042  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/296,014A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1781-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1083 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-296-014A-2

Query Match 62.7%; Score 32; DB 1; Length 1083;  
Best Local Similarity 54.5%; Pred. No. 3.8e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADLSFKSOEL 11  
:||: ||:|  
DB 468 SDLNGIKSEEL 478

Search completed: March 13, 2003, 18:53:51  
Job time : 10.8 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:45:22 ; Search time 22.8 seconds  
(without alignments)  
99.409 Million cell updates/sec

Title: US-10-050-200-4  
Perfect score: 51  
Sequence: 1 ADLSPKSOEL 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_21:\*  
2: sp.archaea:\*  
3: sp.bacteria:\*  
4: sp.fungi:\*  
5: sp.human:\*  
6: sp.invertebrate:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	72.5	305	5 P90872	P90872 caenorhabd
2	37	72.5	4131	5 Q19542	Q19542 caenorhabd
3	36	70.6	173	16 Q9PJR6	Q9PJR6 chlamydia m
4	36	70.6	289	2 Q9F0B9	Q9F0B9 lacobacill
5	36	70.6	293	16 Q9PR79	Q9PR79 ureaplasma
6	36	70.6	379	16 Q982F6	Q982F6 rhizobium
7	36	70.6	1677	5 Q00805	Q00805 leishmania
8	35	68.6	183	5 Q9V0B0	Q9V0B0 drosophila
9	35	68.6	467	17 Q8TT02	Q8TT02 methanosarc
10	35	68.6	1310	5 Q9VYF8	Q9VYF8 drosophila
11	35	68.6	2204	12 Q9WMH6	Q9WMH6 newcastle d
12	34	66.7	137	10 Q9FVY2	Q9FVY2 oryza sativ
13	34	66.7	170	16 Q84483	Q84483 chlamydia t
14	34	66.7	230	16 Q31711	Q31711 bacillus su
15	34	66.7	588	5 Q965R1	Q965R1 caenorhabd
16	34	66.7	630	5 Q95Y88	Q95Y88 caenorhabd

17	33	64.7	75	9 Q38347	Q38347 lactococcus
18	33	64.7	142	2 Q30718	Q30718 helicobacte
19	33	64.7	142	16 Q25037	Q25037 helicobacte
20	33	64.7	142	16 Q92M10	Q92M10 helicobacte
21	33	64.7	174	16 Q44884	Q44884 borrelia bu
22	33	64.7	193	16 Q8RA22	Q8RA22 thermomane
23	33	64.7	205	16 Q8Y9T5	Q8Y9T5 listeria mo
24	33	64.7	209	10 Q9SKF7	Q9SKF7 arabidopsis
25	33	64.7	263	5 Q9XV47	Q9XV47 caenorhabd
26	33	64.7	265	17 Q972E4	Q972E4 sulfolobus
27	33	64.7	388	17 Q974U5	Q974U5 sulfolobus
28	33	64.7	458	16 Q9A409	Q9A409 caulobacter
29	33	64.7	507	4 Q96K90	Q96K90 homo sapien
30	33	64.7	694	4 Q968Y4	Q968Y4 homo sapien
31	33	64.7	694	4 Q96006	Q96006 homo sapien
32	33	64.7	701	16 Q92EA7	Q92EA7 listeria in
33	33	64.7	712	5 Q9U2S0	Q9U2S0 caenorhabd
34	33	64.7	1320	4 Q96KE5	Q96KE5 homo sapien
35	33	64.7	1526	11 Q55078	Q55078 homo sapien
36	33	64.7	1526	11 Q55079	Q55079 cricetus
37	33	64.7	1812	5 Q97275	Q97275 plasmodium
38	32	62.7	81	10 Q858N3	Q858N3 arabidopsis
39	32	62.7	108	3 Q9P6S2	Q9P6S2 schizosacch
40	32	62.7	108	3 P78938	P78938 schizosacch
41	32	62.7	149	5 Q01355	Q01355 halocynthia
42	32	62.7	149	5 Q8WSP7	Q8WSP7 halocynthia
43	32	62.7	149	5 Q8WSP5	Q8WSP5 halocynthia
44	32	62.7	150	5 Q8WSP6	Q8WSP6 halocynthia
45	32	62.7	176	2 Q9AJB8	Q9AJB8 anabaena va

## ALIGNMENTS

RESULT 1  
P90872 PRELIMINARY: PRT: 305 AA.  
AC P90872; P90823;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
DE F39H2.3 protein.  
GN F39H2.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoidea; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA McElay K.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RX MEDLINE=94150718; PubMed=7906398;  
RN [2]  
SEQUENCE FROM N.A.  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
RA Watson A., Watson K., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
RA Weller J., Weller J., Weller J., Weller J., Weller J., Weller J.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RL Nature 368:32-38(1994).  
RN [3]  
SEQUENCE FROM N.A.  
RP Kershaw J.;  
RA Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z81098; CAB03184.1; -.  
DR EMBL; Z81080; CAB03184.1; JOINED.

DR EMBL: 281080; CAB03089.1; JOINED.  
 DR EMBL: 281098; CAB03089.1; JOINED.  
 SQ SEQUENCE 305 AA; 34360 MW; 60D1FAE0AA9666A6 CRC64;

Query Match  
 Best Local Similarity 72.5%; Score 37; DB 5; Length 305;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSSFSKSOEL 11  
 |||||  
 DB 265 LSSFSKSOEL 273

## RESULT 2

OY 019542 PRELIMINARY; PRT; 4131 AA.  
 AC 019542;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 GN F18C12.1 protein.  
 GN F18C12.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B.R.;  
 RU Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RL Science 282:2012-2018(1998).  
 DR EMBL: 275536; CAA99830.1; -.  
 DR HSSP: P02632; ICB1.  
 DR InterPro: IPR004273; Dyrneln\_heavy.  
 DR Pfam: PF03028; Dyrneln\_heavy.1.  
 SQ SEQUENCE 4131 AA; 469655 MW; FD5727CA96A7B2BB CRC64;

Query Match  
 Best Local Similarity 72.5%; Score 37; DB 5; Length 4131;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSSFSKSOEL 11  
 :|||  
 DB 1515 SDSLSSFSKSOEL 1525

## RESULT 3

OY 09PR6 PRELIMINARY; PRT; 173 AA.  
 AC 09PR6;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE DNA methyltransferase.  
 GN TC0762.  
 OS Chlamydia muridarum.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=83560;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MOPN / NIGG;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heldberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
 RA Gwinn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,  
 RA Eisen J., Fraser C.M.;

RT "Genome sequences of Chlamydia trachomatis MOPN and Chlamydia  
 RT pneumoniae AR39."  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 DR EMBL: AE002344; AAF73604.1; -.  
 DR TIGR: TC0762; -.  
 DR InterPro: IPR001497; Methyltransf\_1.  
 DR Pfam: PF01035; Methyltransf\_1; 1.  
 DR TIGRPFAMs: TIGR00589; OG1.1.  
 DR PROSITE: PS00374; MGMT; UNKNOWN\_1.  
 KW Complete proteome.

SQ SEQUENCE 173 AA; 19209 MW; 206DBF41EE94962 CRC64;

Query Match  
 Best Local Similarity 70.6%; Score 36; DB 16; Length 173;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 DLSSFSKSOEL 11  
 |||||  
 DB 93 DLSSFSKSOEL 102

## RESULT 4

OY 09F0B9 PRELIMINARY; PRT; 289 AA.  
 AC 09F0B9;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE EpsI.  
 GN EPSI.  
 OS Lactobacillus delbrueckii (subsp. bulgaricus).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Lactobacillaceae; Lactobacillus.  
 OX NCBI\_TaxID=1585;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LF15;  
 RA Lamothe G.T.;  
 RT "EPS cluster of Lactobacillus delbrueckii subsp. bulgaricus LF15."  
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF267127; AAG44713.1; -.  
 SQ SEQUENCE 289 AA; 34616 MW; CAP418CC19A0F314 CRC64;

Query Match  
 Best Local Similarity 70.6%; Score 36; DB 2; Length 289;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 DLSSFSKSOEL 10  
 |||||  
 DB 225 DLSSFSKSOEL 233

## RESULT 5

OY 09PR79 PRELIMINARY; PRT; 293 AA.  
 AC 09PR79;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Phosphotransacetylase.  
 GN PTA OR U00066.  
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscates;  
 OC Mycoplasmataceae; Ureaplasma.  
 OX NCBI\_TaxID=134821;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SEROVAR 3;  
 RX MEDLINE=20500219; PubMed=11048724;  
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
 RA Cassell G.H.;  
 RT "The complete sequence of the mucosal pathogen Ureaplasma  
 RT urealyticum.";



RL Nature 407:757-762(2000).  
 DR EMBL: AEO02106; AAF30471.1; -  
 DR InterPro: IPR002505; PTA\_PTB.  
 DR Pfam: PF01515; PTA\_PTB; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 293 AA; 33369 MW; 7D51720680FBA4E2 CRC64;

Query Match  
 Best Local Similarity 70.6%; Score 36; DB 16; Length 293;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 DLSSFKSOEL 11  
 |||||: 1:  
 Db 48 DLSSFKTPEI 57

RESULT 6  
 O982F6 PRELIMINARY; PRT; 379 AA.

AC O982F6:  
 DT 01-OCT-2001 (TReMBLrel. 18, Created)  
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, last annotation update)  
 DE 8-amino-7-oxononanoate synthase.  
 GN MLI9099.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI\_TaxID=381;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-MAF3030399;  
 RX MEDLINE-21082930; PubMed-11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,

RA Takeuchi C., Yamada M., Tabata S.,

RT \*Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: APO03015; BAB54503.1; -

DR InterPro: IPR004839; AminoTransf1/2.

KW Pfam: PF00155; aminotran\_1\_2; 1.

KM Plasmid; Complete proteome.

SQ SEQUENCE 379 AA; 40183 MW; 500B3A09095DD314 CRC64;

Query Match  
 Best Local Similarity 70.6%; Score 36; DB 16; Length 379;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADLSSFKSOEL 11  
 |||||: 1:  
 Db 289 ADLVSAQEL 299

RESULT 7  
 O00805 PRELIMINARY; PRT; 1677 AA.

AC O00805:  
 DT 01-JUL-1997 (TReMBLrel. 04, Created)

DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)

DE P-glycoprotein E.

OS Leishmania tropica.

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI\_TaxID=5666;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-LRC-L39;

RA Lafuente E., Castanos S., Gammaro F.;

RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U55381; AAB51191.1; -  
 DR HSSP: P13569; 1MBD.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR001140; ABCtransp1TM.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR InterPro: IPR001865; Ribosomal\_S2.  
 DR InterPro: IPR000504; RNA\_rec\_mol.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC\_transport; 2.  
 DR SMART: SM00382; AAA; 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; UNKNOWN\_1.  
 DR PROSITE: PS00662; RIBOSOMAL\_S2\_1; UNKNOWN\_1.  
 DR PROSITE: PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1677 AA; 183007 MW; C49208921B71563A CRC64;

Query Match  
 Best Local Similarity 70.6%; Score 36; DB 5; Length 1677;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSSFKSOEL 11  
 |||||: 1:  
 Db 669 ADLSTVAQEL 679

RESULT 8  
 O9YU80 PRELIMINARY; PRT; 183 AA.

AC O9YU80:  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TReMBLrel. 14, last annotation update)

DE CG17362 protein.

GN CG17362.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RX MEDLINE-20196006; PubMed-10731132;

RA Adams M.D., Cealiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

Adair J.F., Agdayani A., An H.-D., Andrews-Plankkoch C., Baldwin D.,

Baillet R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,

Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchem K.A.,

Kimmei B.E., Kodira C.D., Kraft C., Krevitz S., Kulp D., Lai Z.,

Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

Liu X., Matvei B., McIntosh T.C., McLeod M.P., McPherson D.,

Merkulov G., Mitsuhashi N.V., Moberly C., Morris J., Moshrefi A.,

Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003536; AAF49779.1; -  
 DR Flybase: FBgn0036393; CG17362.  
 SQ SEQUENCE 183 AA; 19807 MW; BB6263A75679B9A2 CRC64;

Query Match 68.6%; Score 35; DB 5; Length 183;  
 Best Local Similarity 70.0%; Pred. No. 32;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 ADLSPKSOEL 11  
 DB 17 ELSPKSOEL 26

RESULT 9  
 O8TT02 PRELIMINARY: PRT; 467 AA.  
 AC O8TT02;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical protein MA0641.  
 CN MA0641.  
 OS Methanosarcina acetivorans.  
 OC Archaea: Euryarchaeota: Methanococci: Methanosarcinales;  
 CC Methanosarcinaceae: Methanosarcina.  
 OX NCBI\_TaxID=2214;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
 RX MEDLINE=21929760; PubMed=11932238;  
 RA Galagan J.E., Nudbaum C., Roy A., Endrizzi M.G., Macdonald P.,  
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,  
 RA Linton L., Naylor J., Strange-Thomann N., Deatellano K., Johnson R.,  
 RA Linton L., McMan P., McKernan K., Talamas J., Pittrelli A., Ye W.,  
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,  
 RA Hedderley R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,  
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
 RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,  
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
 RA Metcalf W.W., Birren B.;  
 RT "The genome of *Methanosarcina acetivorans* reveals extensive metabolic  
 RT and physiological diversity.";  
 RL Genome Res. 12:532-542(2002).  
 DR EMBL: AE010726; AAM04083.1; -  
 KW Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 467 AA; 51705 MW; FE40724FB890EBF CRC64;

Query Match 68.6%; Score 35; DB 17; Length 467;  
 Best Local Similarity 72.7%; Pred. No. 81;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ADLSPKSOEL 11  
 DB 424 ADLSPKSOEL 434

RESULT 10  
 O9VVF8 PRELIMINARY: PRT; 1310 AA.  
 AC O9VVF8;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE CG7692 protein.  
 GN CG7692.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta;  
 OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha;  
 OC Ephydroidea: Drosophilidae: Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie J.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Blaise R.G., Champs C., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Bencos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,  
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003524; AAF49353.1; -  
 DR Flybase: FBgn0036714; CG7692.  
 SQ SEQUENCE 1310 AA; 150503 MW; 2AF0754E8452430 CRC64;

Query Match 68.6%; Score 35; DB 5; Length 1310;  
 Best Local Similarity 63.6%; Pred. No. 23e02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADLSPKSOEL 11  
 DB 827 ADLSPKSOEL 837

RESULT 11  
 O9WMH6 PRELIMINARY: PRT; 2204 AA.  
 AC O9WMH6;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Large protein.  
 GN L.  
 OS Newcastle disease virus.  
 OC Viruses: ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.  
OX NCBI\_TaxID=11176;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-20047980; PubMed-10580061;  
RA Roemer-Oberdoerfer A., Mundt E., Wekatsion T., Buchholz U.J.,  
RT Mettenleiter T.C.;  
RT "Generation of recombinant lentogenic Newcastle Disease Virus from  
RT CDNA.";  
RL J. Gen. Virol. 80:2987-2995(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Roemer-Oberdoerfer A.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Y18898; CAB51327.1;  
DR InterPro: IPR001016; Viral\_RNA\_pol\_L.  
DR Pfam: PF00946; Paramyx\_RNA\_pol; 1.  
SQ SEQUENCE 2204 AA; 248500 MW; 5E1F50D0E608EDD CRC64;  
  
Query Match 68.6%; Score 35; DB 12; Length 2204;  
Best Local Similarity 70.0%; Pred. No. 3.9e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 DLSSFKSOEL 11  
DB 1393 DLATFKSYEL 1402  
||:|||||  
  
RESULT 12  
Q9FVY2 PRELIMINARY; PRT; 137 AA.  
AC Q9FVY2;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Hypothetical 14.5 kDa protein.  
GN OSUNB0073N24.18.  
GN Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPONBARE;  
RA Bueli C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,  
RA Hsiao J., Zismann V., Pal G., Bowman C.L., Fujii C.Y., Vanaken S.E.,  
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,  
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;  
RT "Oryza sativa chromosome 10 BAC OSUNB0073N24 genomic sequence";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC078840; AAG13611.1;  
KM Hypothetical protein.  
SQ SEQUENCE 137 AA; 14499 MW; 6B7FF4865F97D205 CRC64;  
  
Query Match 66.7%; Score 34; DB 10; Length 137;  
Best Local Similarity 87.5%; Pred. No. 38;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 LSSFKSOE 10  
DB 3 LSSFKAOE 10  
|||||:|  
  
RESULT 13  
O84483 PRELIMINARY; PRT; 170 AA.  
AC O84483;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Methylated-DNA protein - cysteine methyltransferase.  
GN ADA OR CT477.

OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=D/UW-3/CX;  
RX MEDLINE-9900809; PubMed-9784136;  
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
RA Davis R.W.;  
RT "Genome sequence of an obligate intracellular pathogen of humans:  
RT Chlamydia trachomatis.";  
RL Science 282:754-759(1998).  
DR EMBL: AF001321; AAC68077.1;  
DR InterPro: IPR001497; Methyltransf\_1.  
DR Pfam: PF01035; Methyltransf\_1;  
DR PROSITE: PS00374; MGMT; UNKNOWN.1.  
KM Transferase; Methyltransferase; Complete proteome.  
SQ SEQUENCE 170 AA; 19075 MW; 9D5EBA64813245D6 CRC64;  
  
Query Match 66.7%; Score 34; DB 16; Length 170;  
Best Local Similarity 70.0%; Pred. No. 47;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 DLSSFKSOEL 11  
DB 89 DLSSFKKRL 98  
|||||:|  
  
RESULT 14  
O31711 PRELIMINARY; PRT; 230 AA.  
AC O31711;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE YKMY protein.  
GN YKMY.  
OC Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE-98044033; PubMed-9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertoletto M.G., Bessieres P., Bolojin A., Borchert S.,  
RA Bortiss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
RA Denisot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Eutlian K.D., Erington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Funo S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kuno M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
RA Paro V., Pohl T.M., Portetelle D., Portouk S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scifone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin A., Taccori E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terstra P., Tognoni A.,  
RA Toso V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenger T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the gram-positive bacterium *Bacillus*  
 RT subtilis".;  
 RL Nature 390:249-256(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
 RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE-90368558; PubMed-1697575;  
 RA Hemla H., Palva A., Paulin L., Arvidson S., Palva I.;  
 RT "Secretory S complex of *Bacillus subtilis*: sequence analysis and  
 RT identity to pyruvate dehydrogenase".;  
 RL J. Bacteriol. 172:5052-5063(1990).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE-97144523; PubMed-8990290;  
 RA Henriques A.O., Bryan E.M., Beall B.W., Moran C.P. Jr.;  
 RT "cse60, cse60, and csk22 are new members of mother-cell-specific  
 RT sporulation regulons in *Bacillus subtilis*".;  
 RL J. Bacteriol. 179:389-398(1997).  
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS).  
 DR EMBL: Z69111; CAB13309.1; -;  
 DR EMBL: AF012265; AAC24910.1; -;  
 DR InterPro: IPR003593; AAA\_Atpase.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR Pfam: PF00005; ABC\_tran.1.  
 DR ProDom: PD000006; ABC\_transport.1.  
 DR SMART: SM00382; AAA.1.  
 DR PROSITE: PS00211; ABC\_TRANSPORT.1.  
 DR ATP-binding; Transport; Complete proteome.  
 KM ATP-binding; Transport; Complete proteome.  
 SQ SEQUENCE 230 AA; 25272 MW; 59CD1FC50BD619DD CRC64;

Query Match 66.7%; Score 34; DB 16; Length 230;  
 Best Local Similarity 60.0%; Pred. No. 64;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 LSSFKSOEL 11  
 |||||:|:|  
 DB 68 DISSYKDEEL 77

RESULT 15  
 O965R1 PRELIMINARY; PRT; 588 AA.  
 ID O965R1;  
 AC O965R1;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical protein Y97E10B.1.  
 GN Y97E10B.1.  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;  
 CC Rhabditidae; Peloderinae; *Caenorhabditis*.  
 CX NCBI\_TaxID-6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE-99069613; PubMed-9851916;  
 RA None;  
 RT "Genome sequence of the nematode *C. elegans*: a platform for  
 RT investigating biology. The *C. elegans* Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Harmon G., Wohldmann P.;  
 RT "The sequence of *C. elegans* cosmid Y97E10B.";

RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission".;  
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AC024881; AAK71410.1; -;  
 DR InterPro: IPR002875; DUF23.  
 DR Pfam: PF01697; DUF23.1.  
 DR ProDom: PD004153; DUF23.1.  
 SQ SEQUENCE 588 AA; 68177 MW; C293A2B73884BE3F CRC64;

Query Match 66.7%; Score 34; DB 5; Length 588;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 3 LSSFKSOEL 11  
 |||||:|:|  
 DB 386 LSSFKDEEL 394

Search completed: March 13, 2003, 18:51:56  
 Job time : 25.8 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:40:28 ; Search time 5.8 seconds

(without alignments)  
78.662 Million cell updates/sec

Title: US-10-050-200-4  
Perfect score: 51  
Sequence: 1 ADLSEFSGQEL 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	321	1 FCE2_HUMAN	P06734 homo sapien
2	35	68.6	327	1 VP26_HUMAN	O75436 homo sapien
3	35	68.6	514	1 VS14_TRYBB	P26329 trypanosoma
4	35	68.6	523	1 RPN3_YEAST	P40016 saccharomyc
5	35	68.6	2670	1 YAO3_SCHPO	Q10105 schizosacch
6	34	66.7	402	1 PA11_BOVIN	P13909 bos taurus
7	34	66.7	1046	1 SBCC_LACIA	Q9CFZ0 lactococcus
8	34	66.7	1116	1 LN94_YEAST	P53751 saccharomyc
9	33	64.7	316	1 LKHA_DICDI	P52922 dictyosteli
10	33	64.7	364	1 RECF_AZOVI	P49997 azotobacter
11	33	64.7	575	1 GGT_PSESP	P36267 pseudomonas
12	33	64.7	677	1 YD64_MYCPN	P54417 mycoplasma
13	33	64.7	1526	1 TP2A_CRIGR	P41515 cricetus
14	33	64.7	1526	1 TP2A_RAT	P41516 rattus norv
15	33	64.7	1528	1 TP2A_MOUSE	Q01320 mus musculu
16	33	64.7	1531	1 TP2A_HUMAN	P11388 homo sapien
17	33	64.7	1533	1 TP2A_PIG	O46374 sus scrofa
18	32	62.7	146	1 HBB_GYMAC	P29625 gymnodraco
19	32	62.7	258	1 U179_HSV6U	P52469 human herpe
20	32	62.7	332	1 ETEA_CAREL	O93335 caenorhabdi
21	32	62.7	402	1 PA11_PIG	P79433 sus scrofa
22	32	62.7	582	1 ZF36_HUMAN	P16915 homo sapien
23	32	62.7	790	1 RA16_YEAST	P31244 saccharomyc
24	32	62.7	1019	1 LFC_CARRO	Q26422 carciooscor
25	32	62.7	1019	1 LFC_TACTR	P28175 tachyplesu
26	32	62.7	1153	1 ITAM_MOUSE	P05555 mus musculu
27	32	62.7	2204	1 RRP1_NDVB	P11205 newcastile d
28	31	60.8	180	1 NNO1_BUCAI	P57253 buchiera ap
29	31	60.8	322	1 SP5K_BACSU	P27644 bacillus su
30	31	60.8	330	1 ASPG_HELPJ	O25424 helicobacte
31	31	60.8	332	1 ASPG_HELPJ	O25424 helicobacte
32	31	60.8	382	1 YC44_ODOSI	P19539 odontella s
33	31	60.8	545	1 COPD_YEAST	P43621 saccharomyc

## ALIGNMENTS

RESULT 1	FCF2_HUMAN	STANDARD	PRT	321 AA.
AC	P06734			
DT	01-JAN-1988 (Rel. 06, Created)			
DT	01-JAN-1988 (Rel. 06, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Low affinity immunoglobulin epsilon FC receptor (lymphocyte Ige receptor) (Fc-epsilon-RII) (CD23) (BLAST2) (immunoglobulin E-binding factor).			
GN	FCER2 OR IGE2F.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=87118255; PubMed=2949326;			
RA	Ikuta K., Takami M., Kim C.W., Honjo T., Miyoshi T., Tagaya Y., Kawabe T., Yodoi J.;			
RT	"Human lymphocyte Fc receptor for Ige: sequence homology of its cloned cDNA with animal lectins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 84:819-823(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=87051737; PubMed=2877743;			
RA	Kikutani H., Inui S., Sato R., Barsumian E.L., Owaki H., Yamasaki K., Kaisho T., Uchidaayashi N., Hardy R.R., Hirano T., Tsunasawa S., Sakiyama F., Suenura M., Kishimoto T.;			
RT	"Molecular structure of human lymphocyte receptor for immunoglobulin E.";			
RL	Cell 47:657-665(1986).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=87218454; PubMed=3034567;			
RA	Luudin C., Hofstetter H., Sarfati M., Levy C.A., Suter U., Alalmo D., Klichherr E., Frost H., Delespesse G.;			
RT	"Cloning and Expression of the cDNA coding for a human lymphocyte Ige receptor.";			
RL	EMBO J. 6:109-114(1987).			
RN	[4]			
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.			
RA	MEDLINE=93038513; PubMed=1417742;			
RA	Rose K., Turcatti G., Graber P., Pochon S., Regamey P.-O., Jansen K.U., Magneat E., Audonney N., Bonnefoy J.-Y.;			
RT	"Partial characterization of natural and recombinant human soluble CD23.";			
RL	Biochem. J. 286:819-824(1992).			
RN	[5]			
RP	ALTERNATIVE SPLICING.			
RA	MEDLINE=89028672; PubMed=2972386;			
RA	Yokota A., Kikutani H., Tanaka T., Sato R., Barsumian E.L., Suenura M., Kishimoto T.;			
RT	"Two species of human Fc epsilon receptor II (Fc epsilon RII/CD23): tissue-specific and IL-4-specific regulation of gene expression.";			
RL	Cell 55:611-618(1988).			

34	31	60.8	621	1 GIDA_RICPR	O9290 rickettsia
35	31	60.8	853	1 DIA3_HUMAN	O9NSV4 homo sapien
36	31	60.8	1002	1 TSCC_RAT	P55018 rattus norv
37	31	60.8	1021	1 TSCC_HUMAN	P55017 homo sapien
38	31	60.8	1073	1 YAS5_SCHPO	Q10141 schizosacch
39	31	60.8	1115	1 RET_MOUSE	P35546 mus musculu
40	31	60.8	1171	1 DIA3_MOUSE	O92207 mus musculu
41	31	60.8	1319	1 BCCL_ACERY	P19450 acetobacter
42	31	60.8	1520	1 TOP2_CAREL	Q23670 caenorhabdi
43	31	60.8	1549	1 YIR3_YEAST	P40438 saccharomyc
44	31	60.8	1549	1 YJW2_YEAST	P40890 saccharomyc
45	31	60.8	1612	1 TP2B_MOUSE	O64511 mus musculu

```

RN [6]
RP 3D-STRUCTURE MODELING OF LECTIN DOMAIN.
RX MEDLINE-94191542; PubMed-8142907;
RA Padlan E.A., Helm B.A.;
RT "Modeling of the lectin-homology domains of the human and murine low-
affinity Fc epsilon receptor (Fc epsilon RI/CD23).";
RN Receptor 3:325-341(1993).
RP 3D-STRUCTURE MODELING OF 173-285.
RX MEDLINE-96276216; PubMed-8745401;
RA Bajorath J., Aruffo A.;
RT "Structure-based modeling of the ligand binding domain of the human
cell surface receptor CD23 and comparison of two independently
derived molecular models.";
RL Protein Sci. 5:240-247(1996).
CC -1- FUNCTION: THIS RECEPTOR HAS ESSENTIAL ROLES IN THE REGULATION
OF IGE PRODUCTION AND IN THE DIFFERENTIATION OF B-CELLS (IT IS
A B-CELL-SPECIFIC ANTIGEN).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS
A SOLUBLE EXCRETED FORM.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PTM: N- AND O-GLYCOSYLATED.
CC -1- MISCELLANEOUS: THERE ARE TWO KINDS OF FC RECEPTORS FOR IGE, WHICH
DIFFER IN BOTH STRUCTURE AND FUNCTION: HIGH AFFINITY RECEPTORS ON
BASOPHILS AND MAST CELLS AND LOW AFFINITY RECEPTORS ON LYMPHOCYTES
AND MONOCYTES.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- DATABASE: NAME-PROW: NOTE-CD GUIDE CD23 ENTRY.
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd23.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M15059; AAA52434.1; -
DR EMBL: M14766; AAA52435.1; -
DR EMBL: X04772; CAA28465.1; -
DR EMBL: M23562; AAA52433.1; -
DR PIR: A26067; LNHUER.
DR PIR: A26164; A26164.
DR PIR: A26589; A26589.
DR PIR: A31924; A31924.
DR PIR: J10132; J10132.
DR PDB: 1HL1; 31-JAN-94.
DR PDB: 1KJE; 03-APR-96.
DR Genew: HGNC:3612; FCER2.
DR MIM: 151445; -
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; lectin_c.1.
DR SMART: SM00034; CLECT.1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS0041; C-TYPE_LLECTIN_2; 1.
KW Ige-binding protein; Transmembrane; Glycoprotein; Receptor; B-cell;
Repeat; Lectin; Signal-anchor; Alternative splicing; 3D-structure.
FT CHAIN 1 321
FT CHAIN 150 321
FT TRANSMEM 22 47
FT DOMAIN 1 21
FT DOMAIN 150 321
FT TRANSMEM 22 47
FT DOMAIN 48 321
FT DOMAIN 162 284
FT SITE 149 150
FT REPEAT 69 89
FT REPEAT 90 110
FT REPEAT 111 131
FT DISULFID 160 288
FT DISULFID 163 174
FT DISULFID 191 282
BY SIMILARITY.
BY SIMILARITY.

```

```

FT DISULFID 259 273
FT CARBOHYD 63 63
FT VARSPIC 1 7
FT CONFLICT 269 269
SQ SEQUENCE 321 AA: 36468 MW: F86708C0E6515B87 CRC64:

Query Match
Best Local Similarity 100.0%; Score 51; DB 1; Length 321;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADLSFRSQEL 11
DB 118 ADLSFRSQEL 128

RESULT 2
VP26_HUMAN STANDARD; PRT; 327 AA.
AC 075436;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE VPS26 protein homolog.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID-9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Blood;
RX MEDLINE-98318631; PubMed-9653160;
RA Mao M., Fu G., Wu J.-S., Zhang Q.-H., Zhou J., Kan L.-X., Huang Q.-H.,
He K.-L., Gu B.-W., Han Z.-G., Shen Y., Gu J., Yu Y.-P., Xu S.-H.,
Wang Y.-X., Chan S.-J., Chen Z.,
RT "Identification of genes expressed in human CD34(+) hematopoietic
stem/progenitor cells by expressed sequence tags and efficient full-
length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:8175-8180(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Colon;
RA Renfrew Hart C., Sierra L., Balford R., Barr V.A., Taylor S.I.;
RT "Sorting nexins (SNX) 1 and 2: Interaction domains involved in self-
association and associations with human retromer proteins.";
Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE VPS26 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF054179; AAC39912.1; -
DR EMBL: AF175266; AAF89954.1; -
DR Genew: HGNC:12711; VPS26.
DR MIM: 605506; -
DR InterPro: IPR005377; VPS26.
DR Pfam: PF03643; Vps26; 1.
KW Transport; Protein transport.
SQ SEQUENCE 327 AA: 38024 MW: BD330759ABE9BFA9 CRC64:

Query Match
Best Local Similarity 70.0%; Score 35; DB 1; Length 327;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DLSSFRSQEL 11
DB 283 DLSSFRSQEL 292

```

```

RESULT 3
VS14_TRYBB      STANDARD;      PRT;      514 AA.
ID VS14_TRYBB
AC P63329;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Variant surface glycoprotein ILTAT 1.24 precursor (VSG).
OC Trypanosoma brucei brucei.
OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate MING 209;
RX MEDLINE=92046037; PubMed=1942032;
RA Carrington M., Miller N., Blum M.L., Roditi I., Wiley D.C.,
RT Turner M.J.;
RT "Variant specific glycoprotein of Trypanosoma brucei consists of two
RT domains each having an independently conserved pattern of cysteine
RT residues."
RL J. Mol. Biol. 221:823-835(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 24-381.
RX MEDLINE=93218763; PubMed=8464512;
RA Blum M.L., Down J.A., Gurnett A.M., Carrington M., Turner M.J.,
RA Wiley D.C.;
RT "A structural motif in the variant surface glycoproteins of
RT Trypanosoma brucei."
RL Nature 362:603-609(1993).
CC -1- FUNCTION: VSG FORMS A COAT ON THE SURFACE OF THE PARASITE. THE
CC TYPANOSOME EVADES THE IMMUNE RESPONSE OF THE HOST BY EXPRESSING
CC A SERIES OF ANTIGENICALLY DISTINCT VSGS FROM AN ESTIMATED 1000
CC VSG GENES.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC A SOLUBLE FORM IS RELEASED FROM RUPTURED CELLS BY THE ACTION OF A
CC PI-PLC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X56767; CAA40086.1; -.
DR PIR: S18449; S18449.
DR PDB: 2VSG; 25-NOV-98.
DR InterPro: IPR001812; Trypan_glycop.
DR Pfam: PF00913; Trypan_glycop.1.
KW Glycoprotein; Antigen; Trypanosomiasis; GPI-anchor; Membrane; Signal;
KW 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 491 VARIANT SURFACE GLYCOPROTEIN ILTAT 1.24.
FT PROPEP 492 514 HYDROPHOBIC, REMOVED DURING MATURATION.
FT DISULFID 37 162
FT DLSULFID 143 204
FT CARBOHYD 443 443 N-LINKED (GLCNAC, . . .) (POTENTIAL).
FT LIPID 491 491 GPI-ANCHOR (BY SIMILARITY).
SQ SEQUENCE 514 AA; 55169 MW; E77395A9E9CFD8E CRC64;

Query Match 68.6%; Score 35; DB 1; Length 514;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ADLSSFRSDEL 11
Db 56 SDSLNSFKTEL 66
RESULT 4
RPN3_YEAST

```

```

ID RPN3_YEAST      STANDARD;      PRT;      523 AA.
AC P40016;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 26S proteasome regulatory subunit RPN3.
GN RPN3 OR SUN2 OR YER021W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96242146; PubMed=8668124;
RA Kawamura M., Komiyama K.-I., Takeuchi J., Toh E.A.;
RT "A multicopy suppressor of nin1-1 of the yeast Saccharomyces
RT cerevisiae is a counterpart of the Drosophila melanogaster diphenol
RT oxidase A2 gene, Dox-A2."
RL Mol. Gen. Genet. 251:146-152(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97170075; PubMed=9017604;
RA Komiyama K.-I., Okura N., Kawamura M., Demartino G.N., Slaughter C.A.,
RA Shimbara N., Chung C.H., Fujimuro M., Yokosawa H., Shimizu Y.,
RA Tanahashi N., Tanaka K., Toh E.A.;
RT "Yeast counterparts of subunits 55a and p58 (53) of the human 26S
RT proteasome are encoded by two multicopy suppressors of nin1-1."
RL Mol. Biol. Cell 8:171-187(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berio A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mesedale D., Nakahara K., Namath A., Norgren R., Oeffner P., Oh C.,
RA Petzel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yellon M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Acts as a regulatory subunit of the 26S proteasome which
CC is involved in the ATP-dependent degradation of ubiquitinated
CC proteins.
CC -1- SUBUNIT: THE 26S PROTEASOME IS COMPOSED OF A CORE PROTEASE, KNOWN
CC AS THE 20S PROTEASOME, CAPPED AT ONE OR BOTH ENDS BY THE 19S
CC REGULATORY COMPLEX (RC). THE RC IS COMPOSED OF AT LEAST 18
CC DIFFERENT SUBUNITS IN TWO SUBCOMPLEXES, THE BASE AND THE LID,
CC WHICH FORM THE PORTIONS PROXIMAL AND DISTAL TO THE 20S PROTEOLYTIC
CC CORE, RESPECTIVELY (BY SIMILARITY).
CC -1- DOMAIN: CONTAINS 1 PCI DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D78023; BAA1208.1; -.
DR EMBL: U18778; BAB4554.1; -.
DR SGD: S0000823; RPN3.
DR InterPro: IPR000717; PCI.
DR Pfam: PF01399; PCI.1.
DR SMART: SM00088; PINT; 1.
KW Proteasome.
SQ SEQUENCE 523 AA; 60422 MW; DDDA1645B8DE95BD CRC64;

Query Match 68.6%; Score 35; DB 1; Length 523;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

OY 2 DLSFKSQEL 11  
 DB 59 DLSLRNOEL 68

RESULT 5  
 YAO5-SCHPO STANDARD; PRT; 2670 AA.

AC 010105;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Putative transcriptional activator C18G6.05C (GCN1 homolog).  
 GN SPAC18G6.05C.  
 OC Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 NCBI\_TaxID=4896;  
 RX MEDLINE-21848401; PubMed-11859360;  
 RC STRAIN=972;  
 RP SEQUENCE FROM N.A.  
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbittowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hildbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Wandut R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelure V., Mottler S.,  
 RA Gilbert M., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Dague R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forstberg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Useery D., Barrett B.G., Nurse P.;  
 RT \*The genome sequence of Schizosaccharomyces pombe.\*;  
 RL Nature 415:871-880(2002).  
 CC -1- SIMILARITY: SPRONG TO YEAST GCN1.  
 CC -1- SIMILARITY: CONTAINS 19 HEAT REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: Z68198; CAA92385.1;  
 CC DR InterPro: IPR000357; HEAT\_repeat.  
 CC DR Pfam: PF02985; HEAT; 4.  
 CC DR Hypothetical protein; Translation regulation; Activator; Repeat.  
 CC KW PROSITE: PS50077; HEAT\_REPEAT; 4.  
 CC FT REPEAT 315 352 HEAT 1.  
 CC FT REPEAT 1062 1099 HEAT 2.  
 CC FT REPEAT 1319 1356 HEAT 3.  
 CC FT REPEAT 1439 1476 HEAT 4.  
 CC FT REPEAT 1478 1514 HEAT 5.  
 CC FT REPEAT 1518 1555 HEAT 6.  
 CC FT REPEAT 1557 1593 HEAT 7.  
 CC FT REPEAT 1637 1674 HEAT 8.

FT REPEAT 1676 1713 HEAT 9.  
 FT REPEAT 1714 1751 HEAT 10.  
 FT REPEAT 1755 1792 HEAT 11.  
 FT REPEAT 1793 1830 HEAT 12.  
 FT REPEAT 1898 1939 HEAT 13.  
 FT REPEAT 1941 1977 HEAT 14.  
 FT REPEAT 1982 2019 HEAT 15.  
 FT REPEAT 2020 2055 HEAT 16.  
 FT REPEAT 2057 2090 HEAT 17.  
 FT REPEAT 2319 2361 HEAT 18.  
 FT REPEAT 2379 2416 HEAT 19.  
 SQ SEQUENCE 2670 AA; 297333 MW; 948E9316D56D74C3 CRC64;

Query Match 68.6%; Score 35; DB 1; Length 2670;  
 Best Local Similarity 63.6%; Pred. No. 93;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADLSFKSQEL 11  
 DB 672 ADLSRNOEL 682

RESULT 6  
 PA11\_BOVIN STANDARD; PRT; 402 AA.

AC P13909;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Plasminogen activator inhibitor-1 precursor (PAI-1) (Endothelial  
 DE SERPINE1 OR PAI1.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovine; Bos.  
 NCBI\_TaxID=9913;  
 RX MEDLINE-90067867; PubMed-2587231;  
 RA Mimuro J., Sawley M., Hattori M., Loskutoff D.J.;  
 RT \*CDN for bovine type 1 plasminogen activator inhibitor (PAI-1).\*;  
 RL Nucleic Acids Res. 17:8672-8672(1989).  
 RN [2]  
 RP SEQUENCE OF 24-63.  
 RX MEDLINE-88329072; PubMed-3262060;  
 RA Katagiri K., Okada K., Hattori H., Yano M.;  
 RT \*Bovine endothelial cell plasminogen activator inhibitor.  
 RT Purification and heat activation.\*;  
 RL Eur. J. Biochem. 176:81-87(1988).  
 RN [3]  
 RP SEQUENCE OF 153-235 FROM N.A.  
 RC TISSUE=Adrenal cortex;  
 RX MEDLINE-90338128; PubMed-1696269;  
 RA Pepper M.S., Bellin D., Montessano R., Orci L., Vassalli J.-D.;  
 RT \*Transforming growth factor-beta 1 modulates basic fibroblast growth  
 RT factor-induced proteolytic and angiogenic properties of endothelial  
 RT cells in vitro.\*;  
 RL J. Cell Biol. 111:743-755(1990).  
 CC -1- FUNCTION: THIS INHIBITOR ACTS AS "BATE" FOR TISSUE PLASMINOGEN  
 CC ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH  
 CC TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF  
 CC FIBRINOLYSIS.  
 CC -1- MISCELLANEOUS: PAI1 IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE  
 CC UROKINASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE  
 CC 369(R)-370(M) BOND.  
 CC -1- MISCELLANEOUS: VASCULAR ENDOTHELIAL CELLS MAY BE THE PRIMARY SITE  
 CC OF SYNTHESIS OF PLASMA PAI.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its



CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X16383; CAA34419.1; -  
 DR EMBL: X52906; CAA37094.1; -  
 DR PIR: S01324; S01324.  
 DR PIR: S06745; S06745.  
 DR HSSP: P05121; 9PA1.  
 DR InterPro: IPR000215; Serpin.  
 DR Pfam: PF00079; serpin.1.  
 DR SMART: SMO0093; SERPIN.1.  
 DR PROSITE: PS00284; SERPIN.1.  
 KM Serpin, Serine protease inhibitor; Plasma; Plasminogen activation;  
 KW Glycoprotein; Signal.

FT CHAIN 1 23  
 FT SIGNAL 1 23  
 FT ACT\_SITE 369 370 PLASMINOGEN ACTIVATOR INHIBITOR-1.  
 FT CARBOHYD 232 232 REACTIVE BOND.  
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 50 50 S -> L (IN REF. 2).  
 SQ SEQUENCE 402 AA; 45371 MW; 9053617333C7D130 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 402;  
 Best Local Similarity 70.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ADLSEFSQ 10  
 Db 327 ADLSEFSQ 336

RESULT 7  
 ID SBCC\_LACIA STANDARD; PRT; 1046 AA.  
 AC 09CFZ0;  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Nuclease sbccD subunit C.  
 GN SBCC OR IL1321.  
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
 NC NCBI\_TaxID=1360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IL1403;  
 RX MEDLINE=2125186; Pubmed=11337471;  
 RA Bolotin A., Wincker P., Manger S., Jallou O., Malarne K.,  
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;  
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
 RT lactis ssp. lactis IL1403.";  
 RL Genome Res. 11:731-753(2001).  
 CC -1- FUNCTION: sbccD cleaves DNA hairpin structures. These structures  
 CC can inhibit DNA replication and are intermediates in certain DNA  
 CC recombination reactions. The complex acts as a 3'->5' double  
 CC strand exonuclease that can open hairpins. It also has a 5'  
 CC single-strand endonuclease activity (By similarity).  
 CC -1- SUBUNIT: Heterodimer of sbcc and sbcd (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. SBCC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE006364; AAK05419.1; -  
 DR InterPro: IPR003439; ABC\_transportr.

KW Hydrolase; Nuclease; Exonuclease; Endonuclease; DNA replication;  
 KM DNA recombination; ATP-binding; Coiled coil; Complete proteome.

FT NP\_BIND 34 41 ATP (POTENTIAL).  
 FT DOMAIN 223 239 COILED COIL (POTENTIAL).  
 FT DOMAIN 268 432 COILED COIL (POTENTIAL).  
 FT DOMAIN 468 505 COILED COIL (POTENTIAL).  
 FT DOMAIN 543 867 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 1046 AA; 120199 MW; 8F70D00AC28F691 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 1046;  
 Best Local Similarity 70.0%; Pred. No. 54;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLSEFSQ 10  
 Db 728 ADLSEFSQ 737

RESULT 8  
 ID YN94\_YEAST STANDARD; PRT; 1116 AA.  
 AC P53751;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Putative 125.2 kDa membrane glycoprotein in BIO3-HXT17 intergenic  
 DE region.  
 GN YNR065C OR N3539.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NC NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Duesterhoeft A., Floeth M., Fritz C., Heuss-Neitzel D.,  
 RA Hilbert H., Moestl D.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE PEPI FAMILY OF MEMBRANE GLYCOPROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z71680; CAA96347.1; -  
 DR SGD: S0005348; YNR065C.  
 DR InterPro: IPR002860; GH\_BNR.  
 DR Pfam: PF02012; BNR; 7.  
 KM Hypothetical protein; Transmembrane; Glycoprotein.  
 FT TRANSMEM 934 957 POTENTIAL.  
 FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 846 846 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 985 985 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1116 AA; 125199 MW; C03561878F64DAB90 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 1116;  
 Best Local Similarity 77.8%; Pred. No. 58;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 DLSEFSQ 10  
 Db 1056 DLSEFSQ 1064

RESULT 9  
 ID LKHA\_DICD1 STANDARD; PRT; 316 AA.  
 AC P52922;

```

DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Leukotriene A-4 hydrolase (EC 3.3.2.6) (LTA-4 hydrolase) (Leukotriene
DE A(4) hydrolase) (Fragment).
GN LKHA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Jho E., Kopachuk W.;
RL Submitted (Jul-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Hydrolyzes an epoxide moiety of leukotriene A4 (LTA-4)
CC to form leukotriene B4 (LTB-4). The enzyme also has some peptidase
CC activity (By similarity).
CC -1- CATALYTIC ACTIVITY: (7E,9E,11Z,14Z)-(5S,6S)-5,6-epoxycyclo-
CC 7,9,11,14-tetraenoate + H(2)O = (6Z,8E,10E,14Z)-(5S,12R)-5,12-
CC dihydroxyicoso-6,8,10,14-tetraenoate.
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- PATHWAY: Leukotrienes biosynthesis; third step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U27538; AAA70101.1; -
DR MEROPS: M01.004; -
DR DickeyDb: DD05072; LkHA.
DR InterPro: IPR001930; Ala_peptidase.
DR Pfam: PF01433; Peptidase_M1.1.
DR PROSITE: PS00142; ZINC_PROTEASE; PARTIAL.
KW Multifunctional enzyme; Hydrolase; Leukotriene biosynthesis; Zinc;
KW Metalloprotease.
FT NON_TER 1 1 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 1 1 BY SIMILARITY.
FT ACT_SITE 2 2 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 4 4 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 23 23 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 90 90 PROTON DONOR (POTENTIAL).
SQ SEQUENCE 316 AA; 36594 MW; E8307415B91487A CRC64;

Query Match 64.7%; Score 33; DB 1; Length 316;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 LLSFRRQEL 11
DB 196 DIKSFRTQI 205

RESULT 10
RCF_AZOV1 STANDARD; PRT; 364 AA.
AC P49997;
DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE DNA replication and repair protein recf.
GN RECF.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=354;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=UV;
RX MEDLINE=96009876; PubMed=7557415;
RA Badran H., Venkatesh T.V., Kunimajayyan M., Sharma N., Das H.K.;
RT "Molecular characterization of the Azotobacter vinelandii recf gene.";
RL Gene 162:47-51(1995).
CC -1- FUNCTION: THE RECF PROTEIN IS INVOLVED IN DNA METABOLISM; IT IS
CC REQUIRED FOR DNA REPLICATION AND NORMAL SOS INDUCIBILITY. RECF
CC BINDS PREPOTENTIALLY TO SINGLE-STRANDED, LINEAR DNA. IT ALSO SEEMS
CC TO BIND ATP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- SIMILARITY: BELONGS TO THE RECF FAMILY.
CC -1- CAUTION: This entry seems to be produced by an incorrect sequence
CC that contains at least 14 frameshifts. Do not use it for any
CC phylogenetic purpose.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X86404; CAA60158.1; -
DR InterPro: IPR001238; Recf.
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF02463; SMC_N_1.
DR TIGRfams: TIGR00611; recf_1.
DR PROSITE: PS00617; RECF_1.
DR PROSITE: PS00618; RECF_2; 1.
KW DNA damage; DNA replication; DNA-binding; SOS response; DNA repair;
KW ATP-binding.
FT NP_BIND 43 50 ATP (POTENTIAL).
FT DOMAIN 234 259 ARG-RICH (BASIS).
FT DOMAIN 241 244 POLY-ARG.
SQ SEQUENCE 364 AA; 42016 MW; BFF65FDB3AC8407 CRC64;

Query Match 64.7%; Score 33; DB 1; Length 364;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 LLSFRRQEL 11
DB 64 LLSFRRQEL 72

RESULT 11
GCT_PSESP STANDARD; PRT; 575 AA.
AC P36267;
DT 01-JUN-1994 (rel. 29, Created)
DT 01-JUN-1994 (rel. 29, Last sequence update)
DT 01-NOV-1997 (rel. 35, Last annotation update)
DE Gamma-glutamyltranspeptidase precursor (EC 2.3.2.2).
GN GGT.
OS Pseudomonas sp. (strain A14).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=9305254; PubMed=7765305;
RA Ishiy M., Yamashita M., Niwa M.;
RT "Molecular cloning of the gamma-glutamyltranspeptidase gene from a
RT Pseudomonas strain.";
RL Biotechnol. Prog. 9:323-331(1993).
CC -1- CATALYTIC ACTIVITY: (5-L-glutamyl)-peptide + an amino acid =
CC peptide + 5-L-glutamyl-amino acid.
CC -1- PATHWAY: GGT PLAYS A KEY ROLE IN THE GAMMA-GLUTAMYL CYCLE, A
CC PATHWAY FOR THE SYNTHESIS AND DEGRADATION OF GLUTATHIONE.
CC -1- SUBUNIT: THIS ENZYME CONSISTS OF TWO POLYPEPTIDE CHAINS, WHICH ARE
CC SYNTHESIZED IN PRECURSOR FORM FROM A SINGLE POLYPEPTIDE.
CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).

```

```

CC -1- SIMILARITY: STRONG, TO OTHER GGT ENZYMES AND TO GL-7ACA ACYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S63255; AAC60442.1; -.
CC DR MEROPS: T03.001; -.
CC DR InterPro: IPR000101; Gglutinsptase.
CC DR Pfam: PF01019; G_glu_transpt. 1.
CC DR PRINTS: PR01210; GGTTRANSPTASE.
CC DR TIGRFS: TIGR00066; g_glu_trans, 1.
CC DR PROSITE: PS00462; G_GLU_TRANSPEPTIDASE; 1.
CC DR Transfaser: Acyltransferase; Periplasmic; Zymogen; Signal;
CC KM Glutathione biosynthesis.
CC FT SIGNAL 1 24
CC FT CHAIN 25 376 GAMMA-GLUTAMYLTRANSFERASE LARGE CHAIN.
CC FT CHAIN 377 575 GAMMA-GLUTAMYLTRANSFERASE SMALL CHAIN.
CC FT BINDING 450 450 GAMMA-GLUTAMYL (POTENTIAL).
CC FT SEQUENCE 575 AA; 61301 MW; CF2EB69F02CD0201 CRC64;
CC -----
Query Match 64.7%; Score 33; DB 1; Length 575;
Best Local Similarity 54.5%; Pred. No. 45;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 ADLSPKSOE 11
Db 251 ADDQYKTRRL 261

```

```

RESULT 12
ID YD64_MYCPN STANDARD; PRT; 677 AA.
AC P75417;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN364 (H91orf677).
GN MPN364 OR MP472.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Plirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE MG185 / MG260 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000047; AAB96120.1; -.
CC DR InterPro: IPR004890; Lipoprotein_10.
CC DR InterPro: IPR004984; Lipoprotein_X.
CC DR Pfam: PF03202; Lipoprotein_10; 1.
CC DR Pfam: PF03305; Lipoprotein_X; 1.
CC DR Hypothetical protein; Complete proteome.
CC SEQUENCE 677 AA; 75591 MW; E785B68BD679F04D CRC64;

```

```

Query Match 64.7%; Score 33; DB 1; Length 677;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ADLSPKSOE 10
Db 243 ADLSPKSOE 252

```

```

RESULT 13
ID TP2A_CRIGR STANDARD; PRT; 1526 AA.
AC P41515;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA topoisomerase II, alpha isozyme (EC 5.99.1.3).
GN TP2A OR TOP2 OR TOP-2.
OS Citellus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93131977; PubMed=8380592;
RA Chan V.T., Ng S.W., Eder J.P., Schnipper L.E.;
RT "Molecular cloning and identification of a point mutation in the
RT topoisomerase II cDNA from an etoposide-resistant Chinese hamster
RT ovary cell line."
RL J. Biol. Chem. 268:2160-2165(1993).
CC -1- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC MAKES DOUBLE-STRAND BREAKS.
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC RELAX ONLY NEGATIVE SUPERCOILS.
CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L04607; AAA37023.1; -.
CC DR HSSP: P06786; IBGM.
CC DR InterPro: IPR003594; ATPbind_ATPase.
CC DR InterPro: IPR003957; CBFA_NFYB_topis.
CC DR InterPro: IPR001241; DNA_topoisot.
CC DR InterPro: IPR002205; DNA_topoisotIV.
CC DR Pfam: PF00204; DNA_gyraseB; 1.
CC DR Pfam: PF00521; DNA_topoisotIV; 1.
CC DR Pfam: PF02518; HATPase_C; 1.
CC DR PRINTS: PR00615; CCAATSUBUNITA.
CC DR PRINTS: PR00418; TP12FAMILY.
CC DR ProDom: PD000616; DNA_topoisotII; 1.
CC DR ProDom: PD000742; DNA_topoisotIV; 1.
CC DR SMART: SM00433; TOP2c; 1.
CC DR SMART: SM00434; TOP2c; 1.
CC DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
CC DR IsoMase: Topoisomerase; DNA-binding; ATP-binding; Nuclear protein.
CC FT NP_BIND 160 165
CC FT ACT_SITE 804 804
CC FT VARIANT 493 493
CC SEQUENCE 1526 AA; 173196 MW; 5FB2DBFBFIC02929 CRC64;

```

Query Match 64.7%; Score 33; DB 1; Length 1526;  
 Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLSFRSQEL 11  
 ||| ||| |  
 DB 221 DLSFRSQSL 230

## RESULT 14

TP2A\_RAT STANDARD; PRT; 1526 AA.  
 ID TP2A\_RAT  
 AC P41516;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA topoisomerase II, alpha isozyme (EC 5.99.1.3).  
 GN TOP2A OR TOP2 OR TOP-2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RM SEQUENCE FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE-Testis;  
 RX MEDLINE=93290677; PubMed=8390253;  
 RA Park S.H., Yoon J.H., Kwon Y.D., Park S.D.:  
 RT "Nucleotide sequence analysis of the CDNA for rat DNA topoisomerase  
 II.";  
 RL Blochem. Biophys. Res. Commun. 193:787-793(1993).  
 CC -1- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT  
 CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II  
 CC MAKES DOUBLE-STRAND BREAKS.  
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining  
 CC of double-stranded DNA.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH  
 CC RELAX ONLY NEGATIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES  
 CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; 246372; CAAB6496.1; -;  
 CC EMBL; 219552; GAA79611.1; -;  
 CC EMBL; 229676; -; NOT\_ANNOTATED\_CDS.  
 CC HSSP; P06786; 1BGW.  
 CC InterPro; IPR003594; ATPbind\_ATPase.  
 CC InterPro; IPR003957; CBFA\_NFYB\_topIs.  
 CC InterPro; IPR001241; DNA\_topoisomI.  
 CC InterPro; IPR002205; DNA\_topoisomIV.  
 CC Pfam; PF00204; DNA\_gyraseb.1.  
 CC Pfam; PF00521; DNA\_topoisomIV.1.  
 CC Pfam; PF02518; HATPase\_c.1.  
 CC PRINTS; PR00615; CCAATSUBUNTA.  
 CC PRINTS; PR00418; TP12FAMILY.  
 CC ProDom; PD000616; DNA\_topoisomI.1.  
 CC ProDom; PD000742; DNA\_topoisomIV.1.  
 CC SMART; SM00433; TOP2c; 1.  
 CC SMART; SM00433; TOP2c; 1.  
 CC SMART; SM00433; TOP2c; 1.  
 CC PROSITE; PS00177; TOPOISOMERASE\_II; 1.  
 CC PROSITE; PS00177; TOPOISOMERASE\_II; 1.  
 CC Isoenzyme; Topoisomerase: DNA-binding; ATP-binding; Nuclear protein.  
 FT NP\_BIND 159 164 ATP (POTENTIAL).  
 FT ACT\_SITE 803 803 DNA CLEAVAGE (BY SIMILARITY).  
 SQ SEQUENCE 1526 AA; 173220 MW; A1961ABDB1B050F CRC64;

Query Match 64.7%; Score 33; DB 1; Length 1526;  
 Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLSFRSQEL 11  
 ||| ||| |  
 DB 220 DLSFRSQSL 229

## RESULT 15

TP2A\_MOUSE STANDARD; PRT; 1528 AA.  
 ID TP2A\_MOUSE  
 AC Q01320;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA topoisomerase II, alpha isozyme (EC 5.99.1.3).  
 GN TOP2A OR TOP2 OR TOP-2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RM SEQUENCE FROM N.A.  
 RC MEDLINE=93065194; PubMed=1331984;  
 RA Adachi N., Miyake M., Ikeda H., Kikuchi A.;  
 RT "Characterization of cDNA encoding the mouse DNA topoisomerase II  
 RT that can complement the budding yeast top2 mutation.";  
 RL Nucleic Acids Res. 20:5297-5303(1992).  
 CC [2]  
 CC SEQUENCE OF 1254-1528 FROM N.A.  
 CC TISSUE-Lymphoma;  
 CC MEDLINE=94084643; PubMed=8261398;  
 CC McPherson J., Brown G.A., Goldenberg G.J.;  
 RT "Characterization of a DNA topoisomerase IIalpha gene rearrangement  
 RT in adriamycin-resistant P388 leukemia: expression of a fusion  
 RT messenger RNA transcript encoding topoisomerase IIalpha and the  
 RT retinoid acid receptor alpha locus.";  
 RL Cancer Res. 53:5885-5889(1993).  
 CC -1- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT  
 CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II  
 CC MAKES DOUBLE-STRAND BREAKS.  
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining  
 CC of double-stranded DNA.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH  
 CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES  
 CC RELAX ONLY NEGATIVE SUPERCOILS.  
 CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; D12513; BAA02076.1; -;  
 CC EMBL; U01915; AAC52135.1; -;  
 CC PIR; J50703; J50703.  
 CC PIR; S35483; S35483.  
 CC HSSP; P06786; 1BGW.  
 CC MGD; MGI:98790; Top2a.  
 CC InterPro; IPR003594; ATPbind\_ATPase.  
 CC InterPro; IPR003957; CBFA\_NFYB\_topIs.  
 CC InterPro; IPR001241; DNA\_topoisomI.  
 CC InterPro; IPR002205; DNA\_topoisomIV.  
 CC Pfam; PF00204; DNA\_gyraseb.1.  
 CC Pfam; PF00521; DNA\_topoisomIV.1.  
 CC Pfam; PF02518; HATPase\_c.1.

DR PRINTS; PR00615; CCAATSUBUNTA.  
DR PRINTS; PR00418; TP1ZFAMILY.  
DR Prodom; PD000616; DNA\_topoisoi1: 1.  
DR Prodom; PD000742; DNA\_topoisoi1: 1.  
DR SMART; SM00433; TOP2c; 1.  
DR SMART; SM00434; TOP4c; 1.  
DR PROSITE; PS00177; TOPOISOMERASE\_II; 1.  
KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Nuclear protein.  
FT NP\_BIND 160 165 ATP (POTENTIAL).  
FT ACT\_SITE 804 804 DNA CLEAVAGE (BY SIMILARITY).  
SQ SEQUENCE 1528 AA; 172876 MW; 9061778DCAC6C756A CRC64;

Query Match 64.7%; Score 33; DB 1; Length 1528;  
Best local Similarity 70.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLSSFKSQEL 11  
||| ||| |  
Db 221 DLSSFKMQSL 230

Search completed: March 13, 2003, 18:49:53  
Job time : 8.8 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:17 ; Search time 10.8 Seconds  
(without alignments)  
97.915 Million cell updates/sec

Title: US-10-050-200-4

Perfect score: 51

Sequence: 1 ADLSSFKSQEL 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	321	1 LNHR	IgE Fc receptor II
2	37	72.5	305	2 T22009	hypothetical prote
3	37	72.5	4131	2 T21085	hypothetical prote
4	36	70.6	293	2 C82938	phosphotransacetyl
5	36	70.6	1677	2 T18344	p-glycoprotein E-
6	35	68.6	514	2 S18449	variant surface gl
7	35	68.6	533	2 S50479	26S proteasome reg
8	35	68.6	2670	2 T37919	GCN1 homolog - fls
9	34	66.7	170	2 B71510	probable methyltra
10	34	66.7	230	2 D69858	ABC transporter (A
11	34	66.7	270	2 T16227	hypothetical prote
12	34	66.7	402	1 S06745	plasmidogen activa
13	34	66.7	1046	2 A86790	ATP-dependent dSDN
14	34	66.7	1116	2 S63397	probable membrane
15	33	64.7	142	2 H64551	hypothetical prote
16	33	64.7	142	2 B71955	hypothetical prote
17	33	64.7	174	2 G70220	exported protein A
18	33	64.7	205	2 AG1129	hypothetical prote
19	33	64.7	209	2 H84504	En/Spm-like transp
20	33	64.7	263	2 T21619	hypothetical prote
21	33	64.7	265	2 A99249	hypothetical prote
22	33	64.7	364	2 JC4249	recf protein - Azo
23	33	64.7	458	2 H87624	peptidase, M23/M37
24	33	64.7	677	2 S73798	MG260 homolog H91
25	33	64.7	701	2 A11501	internalin protein
26	33	64.7	712	2 T31523	hypothetical prote
27	33	64.7	1526	2 JN0598	DNA topoisomerase
28	33	64.7	1526	2 A44406	DNA topoisomerase
29	33	64.7	1528	2 JS0703	DNA topoisomerase

30	33	64.7	1530	2 A40493	DNA topoisomerase
31	32	62.7	142	2 JC5611	tropomyosin I alpha
32	32	62.7	146	2 S20271	hemoglobin beta ch
33	32	62.7	199	2 T44012	hypothetical prote
34	32	62.7	253	2 C81406	probable hemeolys
35	32	62.7	258	2 T44199	hypothetical prote
36	32	62.7	274	2 T48591	ubiquinol-cytochro
37	32	62.7	312	2 F96758	hypothetical prote
38	32	62.7	316	2 D87839	protein F27D4.111
39	32	62.7	316	2 T21451	hypothetical prote
40	32	62.7	337	2 T48341	hypothetical prote
41	32	62.7	385	2 A71146	hypothetical prote
42	32	62.7	393	2 T20403	hypothetical prote
43	32	62.7	457	2 S66270	kynurenine-oxogluc
44	32	62.7	475	2 E84745	probable splicing
45	32	62.7	499	2 S43324	zeta-carotene desa

#### ALIGNMENTS

RESULT 1  
LNHR  
Ige Fc receptor II, low-affinity [validated] - human  
N/Alternate names: Blast-2; CD23; Fc-epsilon-R1; Lymphocyte Ige receptor  
N/Contents: Ige Fc receptor II, splice form a; Ige Fc receptor II, splice form a'; I;  
C/Species: Homo sapiens (man)  
C/Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 15-Sep-2000  
C/Accession: A26067; S03279; S39442; S39443; A26164; A26589; A31924; J0132; S29107  
R/Rikutan, H.; Inui, S.; Sato, R.; Barsumian, E.L.; Owaki, H.; Yamasaki, K.; Kalsho,  
Cell 47, 657-665, 1986  
A/Title: Molecular structure of human lymphocyte receptor for immunoglobulin E.  
A/Reference number: A26067; MUID:8705173; PMID:2877743  
A/Accession: A26067  
A/Molecule type: mRNA  
A/Residues: 1-321 <K1X>  
A/Cross-references: GB:W14766; NID:g182449; PIDN:AA52435.1; PID:g182450  
A/Experimental source: EBV-transformed B lymphoblastoid cells RPMI-8866  
R/Suter, U.; Bastos, R.; Hofstetter, H.  
Nucleic Acids Res. 15, 7295-7308, 1987  
A/Title: Molecular structure of the gene and the 5'-flanking region of the human lymph  
A/Reference number: S03279; MUID:88015596; PMID:2958779  
A/Accession: S03279  
A/Status: nucleic acid sequence not shown: not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 157-284 <SUT>  
A/Cross-references: GB:X06049; NID:g31316  
A/Note: all exon sequences were determined but the complete sequence is not shown  
R/Matsu, M.; Nunez, R.; Sachl, Y.; Lynch, R.G.; Yodoi, J.  
FEBS Lett. 335, 51-56, 1993  
A/Title: Alternative transcripts of the human CD23/Fc-epsilon-R1. A possible novel  
A/Reference number: S39442; MUID:94063078; PMID:8243664  
A/Accession: S39442  
A/Molecule type: DNA  
A/Residues: 1-7, 'D', '47-50 <MAS1>  
A/Experimental source: splice form a'  
A/Accession: S39443  
A/Molecule type: DNA  
A/Residues: 'MNPSOD', '47-50 <MAS2>  
A/Experimental source: splice form b'  
R/Ludin, C.; Hofstetter, H.; Sarfati, M.; Levy, C.A.; Suter, U.; Alaimo, D.; Klichei  
EMBO J. 6, 109-114, 1987  
A/Title: Cloning and expression of the cDNA coding for a human lymphocyte Ige receptc  
A/Reference number: A26164; MUID:87218454; PMID:3034567  
A/Accession: A26164  
A/Molecule type: mRNA  
A/Residues: 1-268, 'T', '270-321 <LUD>  
A/Cross-references: GB:X04772; NID:g34002; PIDN:CAA28465.1; PID:g34003  
A/Note: the codon given for 265-Asn (ACC) is inconsistent with the authors' translati  
A/Note: part of this sequence, including the amino end of soluble forms of the protei  
R/Ikuta, K.; Takami, M.; Kim, C.W.; Honjo, T.; Miyoshi, T.; Tagaya, Y.; Kawabe, T.; y  
Proc. Natl. Acad. Sci. U.S.A. 84, 819-823, 1987  
A/Title: Human lymphocyte Fc receptor for Ige: sequence homology of its cloned cDNA

A:Reference number: A26589; MUID:87118255; PMID:2949326  
 A:Accession: A26589  
 A:Molecule type: mRNA  
 A:Residues: 1-321 <IKU>  
 A:Cross-references: GB:M15059; NID:9182447; PIDN:AAA52434.1; PID:9182448  
 A:Note: part of this sequence, including the amino end of soluble forms of the protein, R:Tokota, A.; Kikutan, H.; Tanaka, T.; Sato, R.; Barshtan, E.L.; Suenawa, M.; Kishimoto Cell 55, 611-618, 1988  
 A:Title: Two species of human Fc-epsilon receptor II (Fc-epsilon-RII/CD23): tissue-specific  
 A:Reference number: A31924; MUID:89028672; PMID:2972386  
 A:Accession: A31924  
 A:Molecule type: mRNA  
 A:Residues: 'MNPSQ', 8-14 <YOK>  
 A:Cross-references: GB:M23562; NID:9182444  
 A:Experimental source: splice form IIB  
 R:Letellier, M.; Sarfati, M.; Delespasse, G.  
 Mol. Immunol. 26, 1105-1112, 1989  
 A:Title: Mechanisms of formation of IgE-binding factors (soluble CD23)-I. Fc epsilon R I  
 A:Reference number: J10132; MUID:90220658; PMID:2534424  
 A:Accession: J10132  
 A:Molecule type: protein  
 A:Residues: 1-321 <LEP>  
 A:Experimental source: lymphoblastoid B cell line  
 R:Rose, K.; Turcatti, G.; Graber, P.; Pochon, S.; Regamey, P.O.; Jansen, K.U.; Magnenat, Blochem. J. 286, 819-824, 1992  
 A:Title: Partial characterization of natural and recombinant human soluble CD23.  
 A:Reference number: S29107; MUID:93038513; PMID:1417742  
 A:Accession: S29107  
 A:Molecule type: protein  
 A:Residues: 152-166;173-179;189-212;230-263;268-306 <ROS>  
 R:Padlan, E.A.; Helm, B.A.  
 submitted to the Brookhaven Protein Data Bank, June 1993  
 A:Reference number: A51791; PDB:1HLI  
 A:Contents: annotation; conformation by theoretical model, residues 173-285  
 R:Bajorath, J.  
 submitted to the Brookhaven Protein Data Bank, November 1995  
 A:Reference number: A65963; PDB:1KJF  
 A:Contents: annotation; conformation by theoretical model, residues 173-285  
 C:Comment: The sequence of the splice form a is shown.  
 C:Comment: This receptor for the Fc portion of IgE is expressed in various hematopoietic cells.  
 C:Comment: Splice form a is expressed constitutively in B-cells; b is expressed in other  
 C:Comment: Soluble IgE-binding factors are produced by proteolytic cleavage of IgE Fc re  
 C:Genetics:  
 A:Gene: GDB:FCER2; FCE2  
 A:Cross-references: GDB:118888; OMIM:151445  
 A:Map position: 19p13.3-19p13.3  
 A:Introns: 8/1; 46/1; 64/1; 85/1; 106/1; 127/1; 157/1; 207/3; 243/2  
 C:Superfamily: IgE receptor II; C-type lectin homology  
 C:Keywords: alternative splicing; B-cell; glycoprotein; immunoglobulin receptor; macroph  
 F:'MNPSQ', 47-321/Product: IgE Fc receptor II, splice form b' #status predicted <SFB>  
 F:'MNPSQ', 8-321/Product: IgE Fc receptor II, splice form b #status predicted <SFB>  
 F:1-321/Product: IgE Fc receptor II, splice form a #status predicted <SFA>  
 F:1-22/Domain: intracellular #status predicted <INT>  
 F:1-7,'D', 47-321/Product: IgE Fc receptor II, splice form a' #status predicted <SFAI>  
 F:14-20/Region: stop-transfer sequence  
 F:22-25/Domain: transmembrane #status predicted <TM>  
 F:46-321/Domain: extracellular #status predicted <EXT>  
 F:64-84/Region: 21-residue repeat  
 F:81-321/Product: soluble IgE-binding factor (37k) #status predicted <IGB>  
 F:85-105/Region: 21-residue repeat  
 F:102-321/Product: soluble IgE-binding factor (33k) #status predicted <IGB>  
 F:106-126/Region: 21-residue repeat  
 F:125-321/Product: soluble IgE-binding factor (29k) #status predicted <IGI>  
 F:148-321/Product: soluble IgE-binding factor (25-27k), long form #status experimental  
 F:150-321/Product: soluble IgE-binding factor (25-27k), short form #status experimental  
 F:163-282/Domain: C-type lectin homology <ICHR>  
 F:63/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:147-148/Cleavage site: Lys-leu (unidentified proteinase) #status experimental  
 F:149-150/Cleavage site: Arg-Met (unidentified proteinase) #status experimental  
 F:191-282,289-273/Disulfide bonds: #status experimental

Query Match 100.0%; Score 51; DB 1; Length 321;

Best Local Similarity 100.0%; Pred. No. 0.014;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADLSFKSOEL 11  
 Db 118 ADLSFKSOEL 128

## RESULT 2

T22009  
 hypothetical protein F39H2.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T22009; T23398  
 R:Kershaw, J.

submitted to the EMBL Data Library, October 1996  
 A:Reference number: Z19501

A:Accession: T22009

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-305 <WIL>

A:Cross-references: EMBL:Z81080; PIDN:CAB03089.1; GSPDB:GN00019; CESP:F39H2.3

A:Experimental source: clone F39H2

R:McKay, K.  
 submitted to the EMBL Data Library, October 1996

A:Reference number: Z19735

A:Accession: T23398

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-305 <WIL>

A:Cross-references: EMBL:Z81098; PIDN:CAB03184.1; GSPDB:GN00019; CESP:F39H2.3

A:Experimental source: clone K07A12

C:Genetics:  
 A:Gene: CESP:F39H2.3

A:Map position: 1

A:Introns: 60/2; 138/1; 190/3; 219/1; 250/3

## Query Match

Best Local Similarity 72.5%; Score 37; DB 2; Length 305;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSSFKSOEL 11  
 Db 265 LSSFKSOEL 273

## RESULT 3

T21085  
 hypothetical protein F18C12.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000  
 C:Accession: T21085

R:Harris, B.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z19371

A:Accession: T21085

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4131 <WIL>

A:Cross-references: EMBL:Z75536; PIDN:CAA99830.1; GSPDB:GN00019; CESP:F18C12.1

A:Experimental source: clone F18C12

C:Genetics:  
 A:Gene: CESP:F18C12.1

A:Map position: 1

A:Introns: 54/3; 112/3; 146/1; 294/1; 381/3; 455/3; 596/3; 716/3; 743/2; 782/3; 820/3  
 09/3; 2303/3; 2483/1; 2529/1; 2612/3; 2687/1; 2734/3; 2799/3; 2841/3; 2891/2; 2932/3;

C:Superfamily: dynein heavy chain, cytosolic

Query Match 72.5%; Score 37; DB 2; Length 4131;  
 Best Local Similarity 72.7%; Pred. No. 1.5e+02;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSFKSOEL 11



Db 1515 SDLSFKSOEL 1525

# RESULT 4

C82938  
Phosphotransacetylase U0066 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: C82938  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
Submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum. Alternate views of a min  
A:Reference number: A82870  
A:Accession: C82938  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-293 <GLA>  
A:Cross-references: GB:AE002106; GB:AF222894; NID:96899011; PIDN:AAF30471.1; GSPDB:GN001  
C:Genetics:  
A:Gene: pta; U0066  
A:Genetic code: SGCS

Query Match 70.6%; Score 36; DB 2; Length 293;  
Best Local Similarity 70.0%; Pred. No. 15;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 2 DLSSFKSOEL 11  
|||||: 1:  
Db 48 DLSSFKTPEI 57

# RESULT 5

T18344  
P-glycoprotein E - Leishmania tropica  
C:Species: Leishmania tropica  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 31-Jan-2000  
C:Accession: T18344  
R:Laflente, E.; Castanys, S.; Gamarro, F.  
Submitted to the EMBL Data Library, April 1996  
A:Reference number: Z18880  
A:Accession: T18344  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1677 <LAF>  
A:Cross-references: EMBL:U055381; NID:91916605; PID:91916606; PIDN:AMB51191.1  
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 70.6%; Score 36; DB 2; Length 1677;  
Best Local Similarity 63.6%; Pred. No. 92;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
OY 1 ADLSFKSOEL 11  
|||||: 1:  
Db 669 ADLSFYAQL 679

# RESULT 6

S18449  
Variant surface glycoprotein ILTat 1.24 - Trypanosoma brucei  
C:Species: Trypanosoma brucei  
C>Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 20-Aug-1999  
C:Accession: S18449  
R:Carlington, M.; Miller, N.; Blum, M.; Roditi, I.; Wiley, D.; Turner, M.  
J. Mol. Biol. 221, 823-835, 1991  
A:Title: Variant specific glycoprotein of Trypanosoma brucei consists of two domains each  
A:Reference number: S18445; MUID:92046037; PMID:1942032  
A:Accession: S18449  
A:Molecule type: mRNA  
A:Residues: 1-514 <CAR>  
A:Cross-references: EMBL:X56767; NID:910453; PIDN:CAA40086.1; PID:910454  
C:Superfamily: variant surface glycoprotein

C:Keywords: glycoprotein; membrane protein

Query Match 68.6%; Score 35; DB 2; Length 514;  
Best Local Similarity 63.6%; Pred. No. 43;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSFKSOEL 11  
|||||: 1:  
Db 56 SDLSFKTEL 66

# RESULT 7

S50479  
26S proteasome regulatory particle chain RPN3 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YER021W  
C:Species: Saccharomyces cerevisiae  
C>Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 23-Mar-2001  
C:Accession: S50479  
R:Dietrich, F.S.  
Submitted to the EMBL Data Library, December 1994  
A:Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867, and 1am  
A:Reference number: S50433  
A:Accession: S50479  
A:Molecule type: DNA  
A:Residues: 1-523 <DIE>  
A:Cross-references: EMBL:U18778; NID:9603592; PID:9603613; GSPDB:GN00005; MIPS:YER02  
C:Genetics:  
A:Gene: SGD:RPN3; SUN2; MIPS:YER021W  
A:Cross-references: SGD:S0000823; MIPS:YER021W  
A:Map position: 5R  
C:Superfamily: catechol oxidase A2

Query Match 68.6%; Score 35; DB 2; Length 523;  
Best Local Similarity 70.0%; Pred. No. 44;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 2 DLSSFKSOEL 11  
|||||: 1:  
Db 59 DLSSLRNOEL 68

# RESULT 8

T37919  
GCN1 homolog - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 07-Dec-1999  
C:Accession: T37919  
R:Connor, R.; Churche, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.  
Submitted to the EMBL Data Library, December 1995  
A:Reference number: Z21754  
A:Accession: T37919  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2670 <CON>  
A:Cross-references: EMBL:Z68198; PIDN:CAA92385.1; GSPDB:GN00066; SPDB:SPAC1866.05c  
A:Experimental source: strain 972h; cosmid c1866  
C:Genetics:  
A:Gene: SPDB:SPAC1866.05c  
A:Map position: 1  
A:Introns: 50/1

Query Match 68.6%; Score 35; DB 2; Length 2670;  
Best Local Similarity 63.6%; Pred. No. 2.4e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADLSFKSOEL 11  
|||||: 1:  
Db 672 ADLSFKSOEL 682

# RESULT 9

B71510  
probable methyltransferase - Chlamydia trachomatis (serotype D, strain UM3/Cx)

C:Species: Chlamydia trachomatis  
 C>Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
 C:Accession: B71510  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
 A:Reference number: A71570; MUID:99000809; PMID:9784136  
 A:Accession: B71510  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-170 <ARN>  
 A:Cross-references: GB:AE001321; GB:AE001273; NID:93328906; PIDN:AA068077.1; PID:9332891  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: ada

Query Match  
 Best Local Similarity 70.0%; Score 34; DB 2; Length 170;  
 Pred. No. 22;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DLSSFKSOEL 11  
 11111 : 1  
 Db 89 DLSSFKSKRL 98

RESULT 10  
 D69858  
 ABC transporter (ATP-binding protein) homolog ykny - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 02-Feb-2001  
 C:Accession: D69858  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta, C.; Bron, S.; Bruller, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galliz, A.; Gallier, J.; Hartwood, C.R.; Henault, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumslein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: D69858  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-230 <RUN>  
 A:Cross-references: GB:Z99111; GB:AL009126; NID:92633699; PIDN:CA813309.1; PID:el185026;  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: ykny  
 C:Superfamily: short-chain ATP-binding cassette proteins: ATP-binding cassette homology  
 C:Keywords: ATP: nucleotide binding; P-loop  
 F:21-216/Domain: ATP-binding cassette homology <ABC>  
 F:38-45/Region: nucleotide-binding motif A (P-loop)

Query Match  
 Best Local Similarity 60.0%; Score 34; DB 2; Length 230;  
 Pred. No. 30;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 DLSSFKSOEL 11  
 11111 : 1  
 Db 68 DLSSFKSKRL 77

RESULT 11  
 T16227  
 hypothetical protein F31E8.6 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16227  
 R:Du, Z.  
 Submitted to the EMBL Data Library, April 1996  
 A:Description: The sequence of C. elegans cosmid F31E8.  
 A:Reference number: Z18481  
 A:Accession: T16227  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-270 <DUZ>  
 A:Cross-references: EMBL:U55856; NID:91280154; PID:91280159; PIDN:AAA98024.1; GSPDB:  
 A:Experimental source: strain Bristol N2; clone F31E8  
 C:Genetics:  
 A:Gene: CESP:F31E8.6  
 A:Map position: 2  
 A:Introns: 16/3; 147/2; 208/3; 248/3

Query Match  
 Best Local Similarity 70.0%; Score 34; DB 2; Length 270;  
 Pred. No. 35;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DLSSFKSOEL 11  
 11111 : 1  
 Db 127 DLSSFKSKRL 136

RESULT 12  
 S06745  
 Plasminogen activator inhibitor-1 precursor - bovine  
 N:Alternate names: endothelial-cell plasminogen activator inhibitor; PAI-1  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 28-Feb-1990 #sequence\_revision 22-Apr-1995 #text\_change 18-Jun-1999  
 C:Accession: S06745; A35855; S01324; S10906  
 R:Mimuro, J.; Sawley, M.; Hattori, M.; Luskutoff, D.J.  
 Nucleic Acids Res. 17, 8872, 1989  
 A:Title: CDNA for bovine type I plasminogen activator inhibitor (PAI-1).  
 A:Reference number: S06745; MUID:90067867; PMID:2587231  
 A:Accession: S06745  
 A:Molecule type: mRNA  
 A:Residues: 1-402 <MIM>  
 A:Cross-references: EMBL:X16383; NID:9600; PIDN:CAA34419.1; PID:9601  
 R:Pepper, M.S.; Belin, D.; Montesano, R.; Orci, L.; Vassalli, J.D.  
 J. Cell Biol. 111, 743-755, 1990  
 A:Title: Transforming growth factor-beta 1 modulates basic fibroblast growth factor-1  
 A:Reference number: A35855; MUID:90338128; PMID:1696269  
 A:Accession: A35855  
 A:Molecule type: mRNA  
 A:Residues: 153-235 <PPP>  
 A:Cross-references: EMBL:X52906; NID:9598; PIDN:CAA37094.1; PID:9330005  
 R:Katagiri, K.; Okada, K.; Hattori, H.; Yano, M.  
 Eur. J. Biochem. 176, 81-87, 1988  
 A:Title: Bovine endothelial cell plasminogen activator inhibitor. Purification and he  
 A:Reference number: S01324; MUID:88329072; PMID:3262060  
 A:Accession: S01324  
 A:Molecule type: protein  
 A:Residues: 24-49, 'V', '51-63 <KAT>  
 C:Comment: Three types of PAI have been identified. PAI-1 is an acid-stable glycoprot  
 may be the primary site of synthesis of plasma PAI.  
 C:Comment: This inhibitor acts as "bait" for tissue plasminogen activator, urokinase,  
 y518.  
 C:Comment: Glycosylation is not a prerequisite for inhibitory activity.  
 C:Superfamily: antithrombin III  
 C:Keywords: glycoprotein; serine proteinase inhibitor  
 F:1-23/Domain: signal sequence status predicted <SIG>  
 F:24-402/Product: plasminogen activator inhibitor 1 #status predicted <KAT>  
 F:232,288,352/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:369/inhibitory site: Arg (plasminogen activator) #status predicted

Query Match  
 Best Local Similarity 70.0%; Score 34; DB 1; Length 402;  
 Pred. No. 53;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ADLSSFKSOE 10

Db 327 ADFSSFSOE 336

## RESULT 13

A86790  
ATP-dependent dsDNA exonuclease [imported] - Lactococcus lactis subsp. lactis (strain IL1)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: A86790  
R:holoform, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A:Reference number: A86625; MID:21255166; PMID:11337471  
A:Accession: A86790  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-1046 <STO>  
A:Cross-references: GB:AE005176; PID:g12724301; PIDN:AAK05419.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: SPOC

Query Match 66.7%; Score 34; DB 2; Length 1046;  
Best Local Similarity 70.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSFRSQE 10  
1 1111111  
Db 728 ASLNSFESQE 737

## RESULT 14

S63397  
Probable membrane protein YNR065c - Yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein N3539  
C:Species: Saccharomyces cerevisiae  
C:Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 19-Apr-2002  
C:Accession: S63397  
R:Duysterhoef, A.; Floeth, M.; Fritz, C.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.  
submitted to the Protein Sequence Database, April 1996  
A:Reference number: S62944  
A:Accession: S63397  
A:Molecule type: DNA  
A:Residues: 1-1116 <DUE>  
A:Cross-references: EMBL:Z71680; NID:g1302593; PID:e239847; PID:g1302594; GSPDB:GN00014;  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: MIPS:YNR065c  
A:Cross-references: SGD:S0005348  
A:Map position: 14R  
C:Keywords: transmembrane protein  
F:941-957/Domain: transmembrane #status predicted <TM1>  
F:990-1006/Domain: transmembrane #status predicted <TM2>

Query Match 66.7%; Score 34; DB 2; Length 1116;  
Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 DLSSFRSQE 10  
1 1111111  
Db 1056 DLSSFTSOD 1064

## RESULT 15

H64551  
hypothetical protein HP0256 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: H64551  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalaf, H.G.; Giodok, A.; McKen  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.

Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser,  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MID:97394467; PMID:9252185  
A:Accession: H64551  
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-142 <TOM>  
A:Cross-references: GB:AE00545; GB:AE00511; NID:g2313349; PIDN:AAD07334.1; PID:g231

Query Match 64.7%; Score 33; DB 2; Length 142;  
Best Local Similarity 63.6%; Pred. No. 29;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADLSFRSQEL 11  
1 1111111  
Db 42 AOLSAFRNPET 52

Search completed: March 13, 2003, 18:52:57  
Job time : 12.8 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:40:08 ; Search time 29.2 Seconds  
(without alignments)  
50.197 Million cell updates/sec

Title: US-10-050-200-4

Perfect score: 51

Sequence: 1 ADLSSFKSQEL 11

Scoring table: BLOSUM62

Searched: Gapop 10.0, Gapext 0.5

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

A.GeneSeq-101002:\*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	166	AAW13146	Human soluble CD23
2	51	100.0	167	AAW13148	Human soluble CD23
3	51	100.0	320	AAW13143	Sequence containin
4	51	100.0	321	AAW70105	IGE binding factor
5	51	100.0	321	AAW82073	Recombinant Fc-eps
6	51	100.0	321	AAW81163	Low affinity Fc-ep
7	51	100.0	321	AAW81230	Fc gamma receptor.
8	51	100.0	321	AAW81112	Human IGE binding
9	51	100.0	321	AAW81172	Sequence encoded b
10	51	100.0	321	AAW90120	Human lymphocyte r

11	51	100.0	321	10	AAW90367
12	51	100.0	321	14	AAW42053
13	51	100.0	321	18	AAW13142
14	51	100.0	321	9	AAW82838
15	38	74.5	50	22	AAW86386
16	38	74.5	50	22	AAW95865
17	36	70.6	289	21	AAW54094
18	36	70.6	289	21	AAW43796
19	36	70.6	306	22	AAW06927
20	35	68.6	183	22	AAW68544
21	35	68.6	327	20	AAW41679
22	35	68.6	361	21	AAW53264
23	35	68.6	1310	22	AAW63443
24	34	66.7	170	20	AAW37747
25	34	66.7	1046	23	AAW54658
26	33	64.7	91	22	AAW14361
27	33	64.7	91	22	AAW14597
28	33	64.7	174	17	AAW88503
29	33	64.7	205	23	AAW47747
30	33	64.7	309	22	AAW04910
31	33	64.7	377	23	AAW21827
32	33	64.7	378	23	AAW21814
33	33	64.7	507	22	AAW92617
34	33	64.7	659	21	AAW25107
35	32	62.7	42	21	AAW59808
36	32	62.7	45	21	AAW59807
37	32	62.7	81	21	AAW59806
38	32	62.7	210	21	AAW33820
39	32	62.7	270	21	AAW33242
40	32	62.7	270	21	AAW45071
41	32	62.7	274	21	AAW12783
42	32	62.7	274	21	AAW44283
43	32	62.7	306	21	AAW06929
44	32	62.7	312	22	AAW62614
45	32	62.7	314	21	AAW54114

#### ALIGNMENTS

RESULT 1  
AAW13146  
ID AAW13146 standard; Protein: 166 AA.  
XX  
AC AAW13146;  
XX  
DT 17-JUN-1997 (first entry)  
XX  
DE Human soluble CD23 isoform D.  
XX  
KW Human; soluble; CD23; complementarity determining;  
KW Isoform D; recombinant production.  
XX  
OS Homo sapiens.  
XX  
PN JF09028385-A.  
XX  
PD 04-FEB-1997.  
XX  
PE 27-DEC-1995; 95JP-0341169.  
XX  
PR 28-DEC-1994; 94US-0365103.  
XX  
PA (LYNC/) LYNCH R G.  
PA (MATS/) MATSUI M.  
PA (NUNE/) NUNEZ R M.  
PA (YODO/) YODOI J.  
XX  
DR WPI; 1997-159094/15.  
DR N-PSDB; AAT61959.  
XX  
PT DNA sequence encoding a soluble isoform of CD23 - useful for large  
scale preparation of the protein

```

XX Claim 1: Figure 16; 33pp; Japanese.
PS
CC The cDNA encoding the present sequence, human soluble CD23
CC Isoform D, can be used for the large scale recombinant production
CC of soluble CD23.
CC The specification states that the nucleotide sequences contained
CC in figures 4a-b, 5a-b, 12a-12b and 13a-b are claimed, however this
CC numbering scheme does not relate to the scheme used in the relevant
CC sections of the specification.
XX
SQ Sequence 166 AA;
Query Match 100.0%; Score 51; DB 18; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ADLSSFKSOEL 11
DB 79 ADLSSFKSOEL 89

RESULT 2
AAW13148
ID AAW13148 standard; Protein; 167 AA.
XX
AC AAW13148;
XX
XX 17-JUN-1997 (first entry)
DT
XX Human soluble CD23 isoform C.
DE
XX Human; soluble; CD23; complementarity determining;
KW Isoform C; recombinant production.
XX
OS Homo sapiens.
XX
XX JP09028385-A.
PN
XX
XX 04-FEB-1997.
PD
XX
XX 27-DEC-1995; 95JP-0341169.
PF
XX
XX 28-DEC-1994; 94US-0365103.
PR
XX
XX (LYNC/) LYNGH R G.
PA (MATS/) MATSUI M.
PA (NUNEZ/) NUNEZ R M.
PA (YODO/) YODOI J.
XX
XX WPI: 1997-159094/15.
DR N-PSDB; AAT61961.
XX
XX DNA sequence encoding a soluble isoform of CD23 - useful for large
XX scale preparation of the protein
PT
XX
XX Claim 1: Figure 15; 33pp; Japanese.
PS
XX The cDNA encoding the present sequence, human soluble CD23
XX isoform C, can be used for the large scale recombinant production
XX of soluble CD23.
XX The specification states that the nucleotide sequences contained
XX in figures 4a-b, 5a-b, 12a-12b and 13a-b are claimed, however this
XX numbering scheme does not relate to the scheme used in the relevant
XX sections of the specification.
CC
CC
SQ Sequence 167 AA;
Query Match 100.0%; Score 51; DB 18; Length 167;
Best Local Similarity 100.0%; Pred. No. 0.058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ADLSSFKSOEL 11

```

```

DB 80 ADLSSFKSOEL 90
XXXXXXXXXXXX
RESULT 3
AAW13143
ID AAW13143 standard; Protein; 320 AA.
XX
XX AAW13143;
AC
XX 17-JUN-1997 (first entry)
DT
XX
XX Sequence containing human soluble CD23 isoform B.
DE
XX Human; soluble; CD23; complementarity determining;
KW Isoform B; recombinant production.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..204
FT /note="soluble CD23 isoform B"
FT
XX
XX JP09028385-A.
PN
XX
XX 04-FEB-1997.
PD
XX
XX 27-DEC-1995; 95JP-0341169.
PF
XX
XX 28-DEC-1994; 94US-0365103.
PR
XX
XX (LYNC/) LYNGH R G.
PA (MATS/) MATSUI M.
PA (NUNEZ/) NUNEZ R M.
PA (YODO/) YODOI J.
XX
XX WPI: 1997-159094/15.
DR N-PSDB; AAT61956.
XX
XX DNA sequence encoding a soluble isoform of CD23 - useful for large
XX scale preparation of the protein
PT
XX
XX Claim 1: Pages 18-20; 33pp; Japanese.
PS
XX The cDNA encoding the present sequence, which contains human
XX soluble CD23 isoform B, can be used for the large scale recombinant
XX production of soluble CD23.
XX The specification states that the nucleotide sequences contained
XX in figures 4a-b, 5a-b, 12a-12b and 13a-b are claimed, however this
XX numbering scheme does not relate to the scheme used in the relevant
XX sections of the specification.
CC
CC
SQ Sequence 320 AA;
Query Match 100.0%; Score 51; DB 18; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ADLSSFKSOEL 11
DB 117 ADLSSFKSOEL 127

RESULT 4
AAP70105
ID AAP70105 standard; Protein; 321 AA.
XX
XX AAP70105;
AC
XX 04-FEB-1991 (first entry)
DT
XX
XX Ige binding factor.
DE
XX

```

KW Ige binding factor; Fc epsilon receptor; allergy; RPMI 8866 cells;  
 OS Homo sapiens.  
 XX EP248211-A.  
 PN  
 XX  
 PD 09-DEC-1987.  
 XX  
 PF 29-APR-1987; 87EP-0106265.  
 XX  
 PR 30-APR-1986; 86JP-0101531.  
 PR 04-SEP-1986; 86JP-0209091.  
 PR 05-SEP-1986; 86JP-0210429.  
 XX  
 PA (KURS ) KURARAY CO. LTD.  
 XX  
 PI Honjo T, Yodol J, Takami M;  
 DR WPI: 1987-343202/49.  
 DR N-PSDB; AAN70107.  
 XX  
 PT New Ige binding factor protein - obtd. by culturing cells harbouring a  
 PT class specific Fc receptor on their surface.  
 XX  
 PS Claim 6; page 42-3; 57pp; English.  
 XX  
 CC The protein reacts with anti-Fc epsilon receptor antibody, and is obtd.  
 CC in large quantities. The DNA encoding the protein is obtd. from, eg  
 CC human B cells, human malignant B or T cells, human monocytes or human  
 CC eosinophilic cells. RPMI 8866 cells may also be cultured to produce the  
 CC mRNA. The Ige binding factor and Fc epsilon receptor binds Ige and can  
 CC be used to enhance the prodn. of Ige from Ige producing cells. They can  
 CC be used to treat allergies by eliminating excess blood Ige, and can also  
 XX  
 SQ Sequence 321 AA;  
 XX  
 Query Match 100.0%; Score 51; DB 8; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ADLSSFKSQEL 11  
 DB 118 ADLSSFKSQEL 128  
 XX  
 RESULT 5  
 AAP82073  
 ID AAP82073 standard; protein; 321 AA.  
 XX  
 AC AAP82073;  
 XX  
 DT 17-OCT-1990 (first entry)  
 XX  
 DE Recombinant Fc-epsilon receptor.  
 XX  
 KM Low affinity Fc epsilon receptor; Ige; allergy.  
 XX  
 OS Homo sapiens.  
 XX  
 Key Location/Qualifiers  
 FH 1..21 /note="hydrophilic N-terminal"  
 FT Domain /note="basic cluster involved in  
 FT integration into bilayer"  
 FT 14..20 /label=stop-transfer\_seq  
 FT 22..47 /label-transmembrane region  
 FT /note="hydrophobic residues"  
 FT 63..63 /label-glycosylation\_site  
 FT /note="N-linked"  
 FT 149..150  
 FT Cleavage-site

FT /note="target for trypsin-like proteases"  
 FT Protein 1..149  
 FT /label-membrane\_bound  
 FT Protein 150..150  
 FT /label-soluble  
 FT /note="Claim 7"  
 XX  
 PN EP258489-A.  
 XX  
 PD 09-MAR-1988.  
 XX  
 PF 23-SEP-1986; 86EP-0113073.  
 XX  
 PR 21-AUG-1986; 86EP-0111581.  
 XX  
 PA (CELL-) CELLULAR IMMUNOLOGY.  
 XX  
 PI Kishimoto T, Suemura M, Kikutani H, Barsumian EL;  
 DR WPI: 1988-065437/10.  
 DR P-PSDB; AAP82073.  
 XX  
 PT New human low affinity Fc(epsilon)-receptor and parts - useful  
 PT for treating local or systemic allergic reactions and obtd. by  
 PT recombinant DNA methods.  
 XX  
 PS Claim 4; Page 39; 59pp; English.  
 XX  
 CC DNA encoding the sequence or a part of it can be used to make  
 CC recombinant receptor which is useful for treating Ige allergic  
 CC reactions.  
 XX  
 SQ Sequence 321 AA;  
 XX  
 Query Match 100.0%; Score 51; DB 9; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ADLSSFKSQEL 11  
 DB 118 ADLSSFKSQEL 128  
 XX  
 RESULT 6  
 AAP81163  
 ID AAP81163 standard; protein; 321 AA.  
 XX  
 AC AAP81163;  
 XX  
 DT 15-OCT-1990 (first entry)  
 XX  
 DE Low affinity Fc-epsilon receptor encoded by gene carried by pCEM4.  
 XX  
 KM Low affinity Fc-epsilon receptor; Ige; allergy.  
 XX  
 OS Homo sapiens.  
 XX  
 EP257114-A.  
 PN  
 PD 02-MAR-1988.  
 XX  
 PF 21-AUG-1986; 86EP-0111581.  
 XX  
 PR 21-AUG-1986; 86EP-0111581.  
 XX  
 PA (KISH/) KISHIMOTO T.  
 XX  
 PI Kishimoto T, Suemura M, Kikutani H, Barsumian EL;  
 DR WPI: 1988-057531/09.  
 DR N-PSDB; AAN81485.  
 XX  
 PT New human lower affinity Fc(epsilon)-receptor - useful for treating

PT local and allergic reactions produced by the expression of IGE.  
 XX  
 PS Claim 9; Page 24; 36pp; English.  
 XX  
 XX

SQ Sequence 321 AA;

Query Match 100.0%; Score 51; DB 9; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADLSSFKSQEL 11  
 DB 118 ADLSSFKSQEL 128

RESULT 7

AA081230  
 ID AA081230 standard; protein; 321 AA.

AC AA081230;

DT 25-OCT-1990 (first entry)

DE Fc gamma receptor.

KW Fc epsilon receptor; IGE; allergy.

PN JP63198988-A.

PD 17-AUG-1988.

PF 24-APR-1987; 87JP-0042445.

PK 05-SEP-1986; 86JP-0210429.

PR 24-APR-1987; 87US-0042445.

PA (KURS) KURARAY KK.

DR WPI; 1988-273895/39

DR N-PSDB; AAN81612.

PT New DNA contg. nucleotide sequence encoding Fc epsilon receptor -

PT for obtaining IGE-binding substance used to detect and quantify

PT IGE responsible for allergies.

PS Disclosure; ge 529; 16pp; Japanese.

CC Fc epsilon receptor binds IGE. It is expressed from cell line RPMI8866.

CC It is expressed in large amts. and can be used to remove excess IGE from

CC blood, and to detect and quantify IGE. This would allow the development

CC of therapy and diagnosis of allergy.

SQ Sequence 321 AA;

Query Match 100.0%; Score 51; DB 9; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADLSSFKSQEL 11  
 DB 118 ADLSSFKSQEL 128

RESULT 8  
 AA081112  
 ID AA081112 standard; peptide; 321 AA.  
 AC AA081112;  
 DT 12-DEC-1990 (first entry)  
 DE Human IGE binding factor related polypeptide.

XX Human IGE binding factor; IGE-BFs; allergic diseases; B-cells.  
 KW Homo sapiens.  
 OS  
 XX  
 PN EP254249-A.

PD 27-JAN-1988.

PF 20-JUL-1987; 87EP-0110458.

PR 22-JUL-1986; 86GB-0017862.

PR 07-NOV-1986; 86GB-0026622.

PA (CIBA) CIBA GEIGY AG.

PI Hofstetter H, Klichherr E;

DR WPI; 1988-022917/04.

DR N-PSDB; AAN81437.

PT New IGE binding factor related polypeptide(s) - produced by DNA

PT recombinant methods.

PS Claim 1; Page 36; 48pp; English.

CC The peptide is related to the IGE receptors on human B-cells and,

CC if without the membrane anchoring sequence, to the IGE-BFs

CC of Sartali et al. Immunology 53, 197, 207, 783 (1984).

CC Amino acids between 1-133 starting from the N-terminal may be deleted.

CC This is the membrane-anchoring sequence binding the polypeptide to

CC the cytoplasmic membrane of the B-cells.

CC Alternatively amino acids between 110-130 or 250-321 may also be deleted.

CC The peptide has IGE binding activity and is useful for treating

CC allergic conditions, e.g. as caused by antigens such as pollens,

CC cat danders and house dust mites.

SQ Sequence 321 AA;  
 Query Match 100.0%; Score 51; DB 9; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADLSSFKSQEL 11  
 DB 118 ADLSSFKSQEL 128

RESULT 9  
 AA081172  
 ID AA081172 standard; Protein; 321 AA.  
 AC AA081172;  
 DT 28-MAR-1992 (first entry)

DE Sequence encoded by a gene for the water-soluble part of the

DE human low affinity Fc-epsilon-receptor with the amino acids

DE 150 to 321, comprising the EcoRI insert from pFC-epsilon-R-1.

KW Low affinity Fc-epsilon-receptor; systemic IGE-allergic reaction;

KW therapy.

OS Homo sapiens.  
 PN EP259615-A.  
 PD 16-MAR-1988.  
 PF 06-AUG-1987; 87EP-0111392.  
 PR 11-APR-1987; 87EP-0105425.  
 PR 21-AUG-1986; 86EP-0111581.



PR 23-SEP-1986; 86EP-0113073.  
 PR 05-DEC-1986; 86EP-0116938.  
 PR 06-AUG-1987; 87EP-0111392.  
 XX  
 PA (OSAU ) OSAKA UNIVERSITY.  
 PA (KISH/) KISHIMOTO T.  
 XX  
 PI Kishimoto T, Suemura M, Kikutani H, Barsumian EL;  
 DR WPI, 1988-072124/11.  
 DR N-PSDB; AAN81512.  
 XX  
 PT Recombinant human low affinity Fc(epsilon)- receptor - used for  
 PT the treatment of local and systemic IGE-allergic reactions  
 XX  
 PS Claim 3; Table 3, Page 79-81; 118pp; English.  
 XX  
 CC The inventors claim a human low affinity Fc-epsilon-receptor (FCR)  
 CC with an N-terminal cytoplasmic domain, a C-terminal extracellular  
 CC domain and a mol.wt. of about 46kd. Also claimed are recombinant DNA  
 CC contg. the genetic information for the FCR, vectors contg. the DNA,  
 CC host organisms transformed with the vectors, oligonucleotides coding  
 CC for partial amino acid sequences from FCR, and processes for  
 CC preparing FCR. Preferred embodiments of the present invention are  
 CC pFc-epsilon-R-1 (see AAN81516/P81175) and also AAP81172. At least a  
 CC part of the cDNA SO coding for the AAs 1 to 148 of the Fc-epsilon-R  
 CC is replaced by a eucaryotic cDNA signal SO e.g. an interleukin cDNA  
 CC signal SO e.g. by the BSE-2 signal SO (see PBSF-2-LB-AAN81517/P81176).  
 CC A suitable yeast expression vector is a plasmid contg. the yeast  
 CC ADHI-promoter, a gene coding for the yeast mating factor alpha  
 CC leader peptide (MF alpha leader SO) a multicloning site and the  
 CC yeast ADHI-terminator (see AAN81514/P81173). See also EP-258492.  
 XX  
 SQ Sequence 321 AA;

Query Match 100.0%; Score 51; DB 9; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY. 1 ADLSFRKSQEL 11  
 |||||  
 DB 118 ADLSFRKSQEL 128

RESULT 10  
 AAP90120  
 ID AAP90120 standard; protein; 321 AA.  
 XX  
 AC AAP90120;  
 XX  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Human lymphocyte receptor for immunoglobulin.  
 XX  
 KW Human lymphocyte receptor for immunoglobulin; hypersensitivity; allergy;  
 KW asthma; immunoglobulin E; lymphocyte receptor.  
 XX  
 OS Homo sapiens (Human).  
 XX  
 PN EP324879-A.  
 XX  
 PD 26-JUL-1989.  
 XX  
 PF 20-JAN-1988; 88EP-0100814.  
 XX  
 PR 20-JAN-1988; 88EP-0100814.  
 XX  
 PA (OSAU ) OSAKA UNIVERSITY.  
 XX  
 PI Kishimoto T, Suemura M, Kikutani H, Barsumian EL, Schneider F-J;  
 PI Schwendenwein R, Sommergruber W, Swetly P;  
 XX WPI, 1989-214148/30.

DR N-PSDB; AAN90344.  
 XX  
 PT Soluble recombinant Fc-epsilon receptor  
 PT - used for treatment or prophylaxis of local and  
 PT allergic reactions induced by IGE.  
 XX  
 PS Disclosure; fig 1; 23pp; English.  
 XX  
 CC Whole human lymphocyte receptor for immunoglobulin (see corresp.  
 CC AAN90344). Used to produce highly bioactive water-soluble FCR.  
 CC Water-soluble FCR binds IGE, so it is useful for treating  
 CC hypersensitivity, esp. asthma. Amino acid residue 150 is a  
 CC possible site for trypsin-like proteases.  
 XX  
 SQ Sequence 321 AA;

Query Match 100.0%; Score 51; DB 10; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY. 1 ADLSFRKSQEL 11  
 |||||  
 DB 118 ADLSFRKSQEL 128

RESULT 11  
 AAP90367  
 ID AAP90367 standard; protein; 321 AA.  
 XX  
 AC AAP90367;  
 XX  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Peptide sequence of mutated Fc epsilon receptor gene.  
 XX  
 KW Fc epsilon receptor; cloned gene; IGE; allergy; asthma.  
 XX  
 PN EP321601-A.  
 XX  
 PD 28-JUN-1989.  
 XX  
 PF 22-DEC-1987; 87EP-0119080.  
 XX  
 PR 22-DEC-1987; 87EP-0119080.  
 XX  
 PA (OSAU ) OSAKA UNIVERSITY.  
 XX  
 PI Kishimoto T, Suemura M, Kikutani H, Barsumian E;  
 XX  
 DR WPI, 1989-186101/26.  
 DR N-PSDB; AAN90134, AAN90135, AAN90136.  
 XX  
 PT Cloned genes coding for soluble IGE receptor - comprising modified  
 PT coding sequence of Fc epsilon receptor gene.  
 XX  
 PS Disclosure; fig 1; 19pp; English.  
 XX  
 CC The known peptide sequence of Fc epsilon receptor gene.  
 XX  
 SQ Sequence 321 AA;

Query Match 100.0%; Score 51; DB 10; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY. 1 ADLSFRKSQEL 11  
 |||||  
 DB 118 ADLSFRKSQEL 128

RESULT 12  
 AAR42053  
 ID AAR42053 standard; protein; 321 AA.

```

XX AC AAR42053;
XX DT 18-FEB-1994 (first entry)
XX DE Ige binding factor.
XX KW Ige; binding factor; secretion; expression; signal sequence.
XX OS Synthetic.
XX FH Key
XX FT Region
XX FT 148..321
XX FT /label= Ige-binding_factor
XX PN IL84702-A.
XX PD 08-JUL-1993.
XX PF 03-DEC-1987; 87IL-0084702.
XX PR 03-DEC-1987; 87IL-0084702.
XX PA (CIBA ) CIBA GEIGY AG.
XX DR WPI; 1993-275121/35.
XX PT Recombinant DNA molecule - comprises enhancer and promoter unit
XX PT linked to transcripitive DNA segment and DNA segment comprising
XX PT promoter unit linked to DNA sub-segment
XX PS Disclosure; Page 28; 71pp; English.
XX CC The sequence is of Ige-binding factor, secreted from the plasmid
XX CC PSYG-BF. The plasmid PSYG-BF is a derivative of plasmid PSYG-ER
XX CC which has the sequence coding amino acids 1-147 replaced with a new
XX CC coding region (AAQ55400) allowing secretion of the factor.
XX SQ Sequence 321 AA;

Query Match 100.0%; Score 51; DB 14; Length 321;
Best Local Similarity 100.0%; Pred. NO. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADLSSFKSQEL 11
DB 118 ADLSSFKSQEL 128

RESULT 13
AAW13142
ID AAW13142 standard; Protein; 321 AA.
XX
XX AAW13142;
XX AC
XX DT 17-JUN-1997 (first entry)
XX DE Sequence containing human soluble CD23 isoform A.
XX KW Human; soluble; CD23; complementarity determining;
XX KW isoform A; recombinant production.
XX OS Homo sapiens.
XX FH Key
XX FT Peptide
XX FT 1..205
XX FT /note= "soluble CD23 isoform A"
XX PN JP09028385-A.
XX PD 04-FEB-1997.
XX PP 27-DEC-1995; 95JP-0341169.

```

```

XX PR 28-DEC-1994; 94US-0365103.
XX PA (LYNC/) LYNC R G.
XX PA (MATS/) MATSUI M.
XX PA (NUNE/) NUNEZ R M.
XX PA (YODO/) YODOI J.
XX DR WPI; 1997-159094/15.
XX DR N-PSDB; AAT61955.
XX PT DNA sequence encoding a soluble isoform of CD23 - useful for large
XX PT scale preparation of the protein
XX PS Claim 1; Pages 16-17; 33pp; Japanese.
XX CC The cDNA encoding the present sequence, which contains human
XX CC soluble CD23 isoform A, can be used for the large scale recombinant
XX CC production of soluble CD23.
XX CC The specification states that the nucleotide sequences contained
XX CC in figures 4a-b, 5a-b, 12a-12b and 13a-b are claimed, however this
XX CC numbering scheme does not relate to the scheme used in the relevant
XX CC sections of the specification.
XX SQ Sequence 321 AA;

Query Match 100.0%; Score 51; DB 18; Length 321;
Best Local Similarity 100.0%; Pred. NO. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADLSSFKSQEL 11
DB 118 ADLSSFKSQEL 128

RESULT 14
AAP82839
ID AAP82839 standard; Protein; 336 AA.
XX
XX AAP82839;
XX AC
XX DT 16-NOV-1990 (first entry)
XX DE Human low affinity Fc(epsilon) receptor.
XX KW Fc(epsilon) receptor; low affinity; Ige-allergic reactions.
XX KW EP258492-A.
XX PN 09-MAR-1998.
XX PD 05-DEC-1986; 86EP-0116938.
XX PF 05-DEC-1986; 86EP-0116938.
XX PR 05-DEC-1986; 86EP-0116938.
XX PA (CELL-) CELLULAR IMMUNOLOGY.
XX PI Kishimoto T, Suemura M, Kikutani H, Barsumian EL;
XX DR WPI; 1988-065440/10.
XX DR N-PSDB; AAN82252.
XX PT New human low affinity Fc(epsilon)-receptor and parts - useful for
XX PT treating local or systemic Ige-allergic reactions and obd. by
XX PT recombinant DNA methods.
XX PS Disclosure; ; pp; English.
XX CC This protein is useful in the treatment of local or systemic
XX CC Ige-allergic reactions and is obtd. by recombinant DNA methods. It is
XX CC pref. un-accompanied by associated native glycosylation. Fragments of
XX CC it can also be used. See also AAN82253.

```

SQ Sequence 336 AA;  
Query Match 100.0%; Score 51; DB 9; Length 336;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ADLSSFKSOEL 11  
|||||||  
Db 118 ADLSSFKSOEL 128  
  
RESULT 15  
ABR963396  
ID ABR963396 standard; Protein; 50 AA.  
XX  
AC ABR963396;  
XX  
DT 21-JUN-2002 (first entry)  
XX  
DE Human testicular antigen SEQ ID NO: 1780.  
XX  
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disease; infection; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200155317-A2.  
XX  
PD 02-AUG-2001.  
XX  
PE 17-JAN-2001; 2001WO-US01329.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225477.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227109.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230457.  
PR 06-SEP-2000; 2000US-0230458.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246539.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:49:22 ; Search time 9.4 Seconds  
(without alignments)  
53.938 Million cell updates/sec

Title: US-10-050-200-3  
Perfect score: 50  
Sequence: 1 KEMALRESTS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	281	8 US-08-971-317A-6	Sequence 6, Appl1
2	50	100.0	281	9 US-09-131-237-6	Sequence 6, Appl1
3	50	100.0	281	9 US-10-174-654-10	Sequence 10, Appl1
4	50	100.0	281	10 US-09-802-663-25	Sequence 25, Appl1
5	50	100.0	281	10 US-09-193-663-6	Sequence 6, Appl1
6	50	100.0	281	10 US-09-027-287-6	Sequence 6, Appl1
7	50	100.0	281	10 US-09-252-656B-6	Sequence 6, Appl1
8	50	100.0	281	10 US-09-929-493-6	Sequence 6, Appl1
9	50	100.0	281	10 US-09-927-110-1	Sequence 1, Appl1
10	50	100.0	281	12 US-10-012-452-13	Sequence 13, Appl1
11	41	82.0	278	10 US-09-246-129B-6	Sequence 6, Appl1
12	41	82.0	278	10 US-09-899-059-6	Sequence 6, Appl1
13	41	82.0	279	9 US-08-971-317A-4	Sequence 4, Appl1
14	41	82.0	279	9 US-10-017-910-5	Sequence 5, Appl1
15	41	82.0	279	10 US-09-193-663-4	Sequence 4, Appl1
16	34	68.0	108	10 US-09-791-171-4	Sequence 4, Appl1
17	33	66.0	65	10 US-09-731-872-355	Sequence 355, App
18	33	66.0	130	10 US-09-746-359A-9	Sequence 9, Appl1
19	33	66.0	154	10 US-09-746-359A-8	Sequence 8, Appl1

20	33	66.0	468	10 US-09-765-272-10	Sequence 10, Appl1
21	33	66.0	576	10 US-09-841-132-523	Sequence 523, App
22	33	66.0	576	12 US-10-007-693-73	Sequence 73, Appl1
23	32	64.0	89	10 US-09-867-550-686	Sequence 686, App
24	32	64.0	169	10 US-09-886-241-6	Sequence 6, Appl1
25	32	64.0	510	9 US-09-738-626-6798	Sequence 6798, App
26	32	64.0	898	10 US-09-841-739-11	Sequence 11, Appl1
27	32	64.0	1226	10 US-09-815-242-13646	Sequence 13646, A
28	32	64.0	1501	10 US-09-801-368-346	Sequence 346, App
29	31	62.0	118	10 US-09-771-161A-97	Sequence 97, Appl1
30	31	62.0	284	9 US-09-933-999A-6	Sequence 6, Appl1
31	31	62.0	288	10 US-09-815-242-13365	Sequence 13365, A
32	31	62.0	288	10 US-09-815-242-13618	Sequence 13618, A
33	31	62.0	387	10 US-09-771-161A-188	Sequence 188, App
34	31	62.0	427	10 US-09-765-272-172	Sequence 172, App
35	31	62.0	428	10 US-09-815-242-13284	Sequence 13284, A
36	31	62.0	428	10 US-09-815-242-13624	Sequence 13624, A
37	31	62.0	487	10 US-09-764-870-291	Sequence 291, App
38	31	62.0	713	10 US-09-833-435A-9	Sequence 9, Appl1
39	31	62.0	772	9 US-09-978-295A-264	Sequence 264, App
40	31	62.0	772	9 US-09-978-697-264	Sequence 264, App
41	31	62.0	772	9 US-09-978-192A-264	Sequence 264, App
42	31	62.0	772	9 US-10-066-500-121	Sequence 121, App
43	31	62.0	772	9 US-09-999-832A-264	Sequence 264, App
44	31	62.0	772	9 US-09-978-189-264	Sequence 264, App
45	31	62.0	772	9 US-10-174-590-106	Sequence 106, App

#### ALIGNMENTS

RESULT 1  
US-08-971-317A-6  
US-08-971-317A-6  
Sequence 6, Application US/08971317A  
Patent No. US20010010925A1  
GENERAL INFORMATION:  
APPLICANT: Wiley, Steven R.  
TITLE OF INVENTION: TNF-DELTA LIGAND AND USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/971,317A  
FILING DATE: 17-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Goller, Mimi C  
REGISTRATION NUMBER: 39,046  
REFERENCE/DOCKET NUMBER: 6255, US, 01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (847) 935-7550  
TELEFAX: (847) 938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: No. US20010010925A1e  
US-08-971-317A-6

Query Match 100.0%; Score 50; DB 8; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11  
|||||  
DB 109 KELAIRESTS 119

RESULT 2  
US-09-131-237-6

Sequence 6, Application US/09131237B  
Publication No. US20030027284A1

GENERAL INFORMATION:

APPLICANT: Yu, Guo-Liang

APPLICANT: N1, Jian

APPLICANT: Rosen, Craig A.

TITLE OF INVENTION: Tumor Necrosis Factor Gamma

FILE REFERENCE: P141P3

CURRENT APPLICATION NUMBER: US/09/131,237B

CURRENT FILING DATE: 1998-08-07

EARLIER APPLICATION NUMBER: 60/074,047

EARLIER FILING DATE: 1998-02-09

EARLIER APPLICATION NUMBER: 09/005,020

EARLIER FILING DATE: 1998-01-09

EARLIER APPLICATION NUMBER: 08/461,246

EARLIER FILING DATE: 1995-06-05

EARLIER APPLICATION NUMBER: PCT/US94/12880

EARLIER FILING DATE: 1994-11-07

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 6

LENGTH: 281

TYPE: PRT

ORGANISM: Homo sapiens

US-09-131-237-6

Query Match 100.0%; Score 50; DB 9; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11  
|||||  
DB 109 KELAIRESTS 119

RESULT 3  
US-10-174-654-10

Sequence 10, Application US/10174654

Publication No. US20030044937A1

GENERAL INFORMATION:

APPLICANT: Bienkowski, Michael J

APPLICANT: Mills, Cynthia J

APPLICANT: Jones, David A

TITLE OF INVENTION: TNF-Related Death Ligand

NUMBER OF SEQUENCES: 18

ADDRESS: Pharmacia & Upjohn, Intellectual Property

STREET: 301 Henrietta Street

CITY: Kalamazoo

STATE: MI

COUNTRY: USA

ZIP: 49001

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 Diskette

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/174,654  
FILING DATE: 19-Jun-2002

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Kerber, Lori L.

REGISTRATION NUMBER: 41,113

REFERENCE/DOCKET NUMBER: 6111.N CN1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 616/833-0974

TELEFAX: 616/833-8897

TELEX: 224401

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 281 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Query Match 100.0%; Score 50; DB 9; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11  
|||||  
DB 109 KELAIRESTS 119

RESULT 4  
US-09-802-669-25

Sequence 25, Application US/09802669

Patent No. US2002000490A1

GENERAL INFORMATION:

APPLICANT: Dean, Nicholas M.

APPLICANT: Marcusson, Eric G.

APPLICANT: Wyatt, Jacqueline

APPLICANT: Zhang, Hong

TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling

FILE REFERENCE: ISPH-545

CURRENT APPLICATION NUMBER: US/09/802,669

CURRENT FILING DATE: 2001-03-09

PRIOR APPLICATION NUMBER: US 09/665,615

PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: US 09/290,640

PRIOR FILING DATE: 1999-04-12

NUMBER OF SEQ ID NOS: 180

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 25

LENGTH: 281

TYPE: PRT

ORGANISM: Homo sapiens

US-09-802-669-25

OY 1 KELAIRESTS 11  
|||||  
DB 109 KELAIRESTS 119

RESULT 5  
US-09-193-663-6

Sequence 6, Application US/09193663

Patent No. US20020055624A1

GENERAL INFORMATION:

APPLICANT: Wiley, Steven R.

TITLE OF INVENTION: TNF-DELTA LIGAND AND USES THEREOF

FILE REFERENCE: 6255. US.02

CURRENT APPLICATION NUMBER: US/09/193,663

; CURRENT FILING DATE: 1998-11-17  
; EARLIER APPLICATION NUMBER: 60/065,916  
; EARLIER FILING DATE: 1997-11-17  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-193-663-6

Query Match 100.0%; Score 50; DB 10; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
|||||  
Db 109 KELAELRESTS 119

RESULT 6  
US-09-027-287-6  
; Sequence 6, Application US/09027287A  
; Patent No. US20020064869A1  
; GENERAL INFORMATION:  
; APPLICANT: Edner, Reinhard  
; APPLICANT: Ruben, Steven M.  
; APPLICANT: Ullrich, Stephen  
; TITLE OF INVENTION: Apoptosis Inducing Molecule II  
; FILE REFERENCE: 1488.0650004  
; CURRENT APPLICATION NUMBER: US/09/027,287A  
; EARLIER FILING DATE: 1998-02-20  
; EARLIER APPLICATION NUMBER: US 09/003,886  
; EARLIER FILING DATE: 1998-01-07  
; EARLIER APPLICATION NUMBER: US 08/822,953  
; EARLIER FILING DATE: 1997-03-21  
; EARLIER APPLICATION NUMBER: US 60/030,157  
; EARLIER FILING DATE: 1996-10-31  
; EARLIER APPLICATION NUMBER: US 60/013,923  
; EARLIER FILING DATE: 1996-03-22  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-027-287-6

Query Match 100.0%; Score 50; DB 10; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
|||||  
Db 109 KELAELRESTS 119

RESULT 7  
US-09-252-656B-6  
; Sequence 6, Application US/09252656B  
; Patent No. US20020081647A1  
; GENERAL INFORMATION:  
; APPLICANT: Edner, Reinhard  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: Ruben, Steven M.  
; APPLICANT: Zhang, Jun  
; APPLICANT: Ullrich, Stephen  
; APPLICANT: Zhai, Yifan  
; TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use  
; FILE REFERENCE: 1488.0650006  
; CURRENT APPLICATION NUMBER: US/09/252,656B  
; CURRENT FILING DATE: 1999-02-19

; PRIOR APPLICATION NUMBER: US 60/075,409  
; PRIOR FILING DATE: 1998-02-20  
; PRIOR APPLICATION NUMBER: US 09/027,287  
; PRIOR FILING DATE: 1998-02-20  
; PRIOR APPLICATION NUMBER: US 09/003,886  
; PRIOR FILING DATE: 1998-01-07  
; PRIOR APPLICATION NUMBER: US 08/822,953  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/013,923  
; PRIOR FILING DATE: 1996-03-22  
; PRIOR APPLICATION NUMBER: US 60/030,157  
; PRIOR FILING DATE: 1996-10-31  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-252-656B-6

Query Match 100.0%; Score 50; DB 10; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
|||||  
Db 109 KELAELRESTS 119

RESULT 8  
US-09-929-493-6  
; Sequence 6, Application US/09929493  
; Patent No. US20020115112A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu et al.  
; TITLE OF INVENTION: Neurotrophin-alpha and Neurotrophin-alpha Splice Variant  
; FILE REFERENCE: PF343P4  
; CURRENT APPLICATION NUMBER: US/09/929,493  
; CURRENT FILING DATE: 2001-08-15  
; PRIOR APPLICATION NUMBER: 60/225,628  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/227,008  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: 60/234,338  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/240,806  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/250,020  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-06  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/296,122  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: 60/304,809  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: n equals  
; OTHER INFORMATION: a, t, g, or c  
US-09-929-493-6

Query Match 100.0%; Score 50; DB 10; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11

Db 109 KELAELRESTS 119

RESULT 9

US-09-927-110-1  
; Sequence 1, Application US/09927110  
; Patent No. US20020127233A1  
; GENERAL INFORMATION:  
; APPLICANT: SEABRIGHT CORPORATION LIMITED  
; APPLICANT: ZHU, Bing  
; APPLICANT: CYNADER, Max  
; APPLICANT: PATY, Donald  
; APPLICANT: LYO, Liqing  
; TITLE OF INVENTION: METHOD FOR INHIBITING INFLAMMATION IN IMMUNE PRIVILEGED SITES  
; FILE REFERENCE: MM1240  
; CURRENT APPLICATION NUMBER: US/09/927,110  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: US 60/224,016  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 1  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-927-110-1

Query Match

Best Local Similarity 100.0%; Score 50; DB 10; Length 281;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELAELRESTS 11  
Db 109 KELAELRESTS 119

RESULT 10

US-10-012-452-13  
; Sequence 13, Application US/10012452  
; Patent No. US2002011325A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, et al.  
; TITLE OF INVENTION: VEG1, An Inhibitor of Angiogenesis and Tumor Growth  
; FILE REFERENCE: PF141P5  
; CURRENT APPLICATION NUMBER: US/10/012,452  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 09/559,290  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: PCT/US98/23191  
; PRIOR FILING DATE: 1998-11-02  
; PRIOR APPLICATION NUMBER: 08/963,272  
; PRIOR FILING DATE: 1997-11-03  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-012-452-13

Query Match

Best Local Similarity 100.0%; Score 50; DB 12; Length 281;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELAELRESTS 11  
Db 109 KELAELRESTS 119

RESULT 11  
US-09-246-129B-6

; Sequence 6, Application US/09246129B  
; Patent No. US20020090683A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: NI, Jian  
; APPLICANT: Rosen, Craig A.

; TITLE OF INVENTION: Tumor Necrosis Factor Gamma  
; FILE REFERENCE: PF141P4  
; CURRENT APPLICATION NUMBER: US/09/246,129B  
; CURRENT FILING DATE: 1999-02-08  
; PRIOR APPLICATION NUMBER: 60/074,047  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 09/131,237  
; PRIOR FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: 09/005,020  
; PRIOR FILING DATE: 1998-01-09  
; PRIOR APPLICATION NUMBER: 08/461,246  
; PRIOR FILING DATE: 1995-06-05  
; PRIOR APPLICATION NUMBER: PCT/US94/12880  
; PRIOR FILING DATE: 1994-11-07  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 278  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-246-129B-6

Query Match

Best Local Similarity 82.0%; Score 41; DB 10; Length 278;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELAELRESTS 11  
Db 106 KELAELREFTN 116

RESULT 12

US-09-899-059-6  
; Sequence 6, Application US/09899059  
; Patent No. US20020150534A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: NI, Jian  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Zhang, Jun  
; TITLE OF INVENTION: Tumor Necrosis Factor Gamma  
; FILE REFERENCE: PF141P7  
; CURRENT APPLICATION NUMBER: US/09/899,059  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/278,449  
; PRIOR FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: 60/216,879  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 09/559,290  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/180,908  
; PRIOR FILING DATE: 2000-02-08  
; PRIOR APPLICATION NUMBER: 60/134,067  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: 60/132,227  
; PRIOR FILING DATE: 1999-05-03  
; PRIOR APPLICATION NUMBER: 60/131,963  
; PRIOR FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: 09/246,129  
; PRIOR FILING DATE: 1999-02-08  
; PRIOR APPLICATION NUMBER: 60/074,047  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 09/131,237  
; PRIOR FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: 09/005,020  
; PRIOR FILING DATE: 1998-01-09  
; PRIOR APPLICATION NUMBER: 08/461,246



; PRIOR FILING DATE: 1995-06-05  
; PRIOR APPLICATION NUMBER: PCT/US94/12880  
; PRIOR FILING DATE: 1994-11-07  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO: 6  
; LENGTH: 278  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-899-059-6

Query Match 82.0%; Score 41; DB 10; Length 278;  
Best Local Similarity 81.8%; Pred. No. 2.4;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KEIAELRESTS 11  
Db 106 KEIAELREFTN 116

RESULT 13  
US-08-971-317A-4  
; Sequence 4, Application US/08971317A  
; Patent No. US20010010925A1  
; GENERAL INFORMATION:  
; APPLICANT: Wiley, Steven R.  
; TITLE OF INVENTION: TNF-DELTA LIGAND AND USES  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: IBM Compatible  
; SOFTWARE: FASTSEQ for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/971.317A  
; FILING DATE: 17-NOV-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goller, Mini C  
; REGISTRATION NUMBER: 39,046  
; REFERENCE/DOCKET NUMBER: 6255-US-01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (847) 935-7550  
; TELEFAX: (847) 938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 279 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: NO. US20010010925A1e  
US-08-971-317A-4

Query Match 82.0%; Score 41; DB 8; Length 279;  
Best Local Similarity 81.8%; Pred. No. 2.4;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
OY 1 KEIAELRESTS 11  
Db 107 KEIAELREFTN 117

RESULT 14  
US-10-017-910-5  
; Sequence 5, Application US/10017910  
; Patent No. US20020159970A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi, Yongwon  
; Wong, Brian  
; Josien, Regis  
; Steinman, Ralph

; TITLE OF INVENTION: A PROTEIN BELONGING TO THE TNF SUPERFAMILY  
; INVOLVED IN SIGNAL TRANSDUCTION, NUCLEIC ACIDS ENCODING  
; METHODS OF USE THEREOF

; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/017,910  
; FILING DATE: 14-Dec-2001  
; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/447,035  
; FILING DATE: 1999-11-22  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-200  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 279 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Mus musculus

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-10-017-910-5

Query Match 82.0%; Score 41; DB 9; Length 279;  
Best Local Similarity 81.8%; Pred. No. 2.4;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
OY 1 KEIAELRESTS 11  
Db 107 KEIAELREFTN 117

RESULT 15  
US-09-193-663-4  
; Sequence 4, Application US/09193663  
; Patent No. US20020055624A1  
; GENERAL INFORMATION:  
; APPLICANT: Wiley, Steven R.  
; TITLE OF INVENTION: TNF-DELTA LIGAND AND USES THEREOF  
; FILE REFERENCE: 6255-US-02  
; CURRENT APPLICATION NUMBER: US/09/193,663  
; CURRENT FILING DATE: 1998-11-17

```

; EARLIER APPLICATION NUMBER: 60/065,916
; EARLIER FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-663-4

```

```

Query Match      82.0%; Score 41; DB 10; Length 279;
Best Local Similarity 81.8%; Pred. No. 2.4;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 KELAELRESYS 11
        |||||
Db      107 KELAELREFTN 117

```

Search completed: March 13, 2003, 18:54:45  
 Job time : 10.4 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:52 ; Search time 9.8 Seconds

(without alignments)  
33.026 Million cell updates/sec

Title: US-10-050-200-3

Sequence: 1 KEIARESTS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/6C.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/6ackfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	178	4	US-08-339-214-84
2	50	100.0	179	3	US-08-649-100-9
3	50	100.0	179	4	US-08-339-214-6
4	50	100.0	281	2	US-08-810-453-2
5	50	100.0	281	3	US-08-815-190A-2
6	50	100.0	281	4	US-09-290-640-25
7	50	100.0	281	4	US-09-479-524-3
8	50	100.0	281	4	US-08-339-214-8
9	50	100.0	281	4	US-08-339-214-30
10	50	100.0	281	4	US-09-589-287B-6
11	50	100.0	281	4	US-09-157-864-10
12	50	100.0	281	4	US-09-006-755B-1
13	50	100.0	281	5	PCT-US95-00362-2
14	50	100.0	287	5	US-08-815-190A-16
15	50	100.0	376	3	US-08-751-512-8
16	41	82.0	179	4	US-08-339-214-14
17	41	82.0	179	4	US-08-339-214-14
18	41	82.0	278	4	US-08-339-214-22
19	41	82.0	278	4	US-08-339-214-16
20	41	82.0	279	4	US-08-339-214-26
21	41	82.0	279	4	US-08-339-214-24
22	41	82.0	279	5	PCT-US95-00362-5
23	34	68.0	1045	1	US-08-452-083-2
24	33	66.0	468	4	US-08-961-083-10
25	32	64.0	163	4	US-09-562-737-74
26	32	64.0	323	6	5185254-4
27	31	62.0	98	4	US-09-615-192A-278

28	31	62.0	112	4	US-09-615-192A-276	Sequence 276, App
29	31	62.0	133	4	US-09-615-192A-277	Sequence 277, App
30	31	62.0	256	4	US-09-071-035-300	Sequence 300, App
31	31	62.0	284	4	US-09-071-035-298	Sequence 298, App
32	31	62.0	387	1	US-08-713-828-3	Sequence 3, Appl1
33	31	62.0	387	2	US-08-919-627-3	Sequence 3, Appl1
34	31	62.0	387	2	US-09-096-245-3	Sequence 3, Appl1
35	31	62.0	387	4	US-09-457-040B-30	Sequence 30, Appl1
36	31	62.0	388	1	US-08-713-828-4	Sequence 4, Appl1
37	31	62.0	388	1	US-08-713-828-5	Sequence 4, Appl1
38	31	62.0	388	2	US-08-919-627-4	Sequence 4, Appl1
39	31	62.0	388	2	US-08-919-627-5	Sequence 4, Appl1
40	31	62.0	388	2	US-09-096-245-4	Sequence 4, Appl1
41	31	62.0	388	2	US-09-096-245-5	Sequence 5, Appl1
42	31	62.0	427	4	US-08-961-083-172	Sequence 172, App
43	31	62.0	713	4	US-09-346-237-9	Sequence 23, Appl1
44	31	62.0	1151	3	US-08-836-134-23	Sequence 23, Appl1
45	31	62.0	1151	4	US-09-493-784-23	Sequence 23, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-339-214-84  
Sequence 84, Application US/08339214  
Patent No. 6348334  
GENERAL INFORMATION:  
APPLICANT: Nagata, Shigikazu  
APPLICANT: Suda, Takashi  
APPLICANT: Nakamura, No. 634833410  
TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Birch, Stewart, Kolasch & Birch  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/339, 214  
FILING DATE: 10-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 178 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..178  
OTHER INFORMATION: /note="polypeptide CDI179"  
US-08-339-214-84  
Query Match 100.0% Score 50: DB 4: Length 178:

Best Local Similarity 100.0%; Pred. No. 0.037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 KELALRESTD 11  
DB 7 KELALRESTD 17

RESULT 2  
US-08-649-100-9  
Sequence 9, Application US/08649100  
Patent No. 6114507  
GENERAL INFORMATION:  
APPLICANT: SHIRAKAWA, KAMON  
APPLICANT: MATUSE, TOMOKAZU  
APPLICANT: NAGATA, SHIGEKAZU  
APPLICANT: CO, MAN SUNG  
APPLICANT: VASQUEZ, MAXIMILIANO  
TITLE OF INVENTION: ANTI-FAS LIGAND ANTIBODY AND ASSAY  
TITLE OF INVENTION: METHOD USING THE ANTI-FAS LIGAND ANTIBODY  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH  
STREET: PO BOX 747  
CITY: FALLS CHURCH  
STATE: VA  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/649,100  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURPHY JR, GERALD M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1110-160  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 179 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-649-100-9  
Query Match 100.0%; Score 50; DB 3; Length 179;  
Best Local Similarity 100.0%; Pred. No. 0.037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 KELALRESTD 11  
DB 7 KELALRESTD 17

RESULT 3  
US-08-339-214-6  
Sequence 6, Application US/08339214  
Patent No. 6348334  
GENERAL INFORMATION:  
APPLICANT: Nagata, Shigikazu  
APPLICANT: Suda, Takashi  
APPLICANT: Takahashi, Tomonorio  
APPLICANT: Nakamura, No. 634833410  
TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA  
TITLE OF INVENTION: Encoding the Same

NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/339,214  
FILING DATE: 10-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1110-139P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 179 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-339-214-6  
Query Match 100.0%; Score 50; DB 4; Length 179;  
Best Local Similarity 100.0%; Pred. No. 0.037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 KELALRESTD 11  
DB 7 KELALRESTD 17

RESULT 4  
US-08-810-453-2  
Sequence 2, Application US/08810453  
Patent No. 5858990  
GENERAL INFORMATION:  
APPLICANT: Walsh, Kenneth  
TITLE OF INVENTION: FAS LIGAND COMPOSITIONS FOR TREATMENT OF  
TITLE OF INVENTION: PROLIFERATIVE DISORDERS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: Federal Reserve Plaza, 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02210-2211  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,453  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: 51237/7004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)720-3500  
TELEFAX: (617)720-2441

TELEX: 343  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-810-453-2

Query Match 100.0%; Score 50; DB 2; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.059;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11  
Db 109 KELAIRESTS 119

RESULT 5  
US-08-815-190A-2  
; Sequence 2, Application US/08815190A  
; Patent No. 6046310  
; GENERAL INFORMATION:  
; APPLICANT: Queen, Cary L.  
; APPLICANT: Schneider, William P.  
; APPLICANT: Vasquez, Maximiliano  
; TITLE OF INVENTION: Fas Ligand Fusion Proteins and Their  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834

COMPUTER READABLE FORM:  
; COMPUTER TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/815,190A  
; FILING DATE: 11-MAR-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/614,584  
; FILING DATE: 13-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 011823-006710US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-815-190A-2

Query Match 100.0%; Score 50; DB 3; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.059;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11  
Db 109 KELAIRESTS 119

RESULT 6

US-09-290-640-25  
; Sequence 25, Application US/09290640  
; Patent No. 6204055  
; GENERAL INFORMATION:  
; APPLICANT: Dean, Nicholas M.  
; APPLICANT: Marcussen, Eric G.  
; TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling  
; FILE REFERENCE: ISPH-0351  
; CURRENT APPLICATION NUMBER: US/09/290,640  
; CURRENT FILING DATE: 1999-04-12  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-290-640-25

Query Match 100.0%; Score 50; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.059;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11  
Db 109 KELAIRESTS 119

RESULT 7  
US-09-479-524-3  
; Sequence 3, Application US/09479524  
; Patent No. 6268350  
; GENERAL INFORMATION:  
; APPLICANT: Barbera-Guillien, Emilio  
; TITLE OF INVENTION: Polynucleotides for Inhibiting Metastasis And Tumor Cell Gro  
; FILE REFERENCE: B-37  
; CURRENT FILING DATE: 2000-01-07  
; EARLIER APPLICATION NUMBER: US 09/170,948  
; EARLIER FILING DATE: 1998-10-13  
; EARLIER APPLICATION NUMBER: US 60/062,733  
; EARLIER FILING DATE: 1997-10-23  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: WordPerfect  
; SEQ ID NO 3  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-479-524-3

Query Match 100.0%; Score 50; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.059;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11  
Db 109 KELAIRESTS 119

RESULT 8  
US-08-339-214-8  
; Sequence 8, Application US/08339214  
; Patent No. 6348334  
; GENERAL INFORMATION:  
; APPLICANT: Nagata, Shigikazu  
; APPLICANT: Suda, Takashi  
; APPLICANT: Takahashi, Tomoniro  
; APPLICANT: Nakamura, No. 634833410  
; TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: P.O. Box 747

CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/339,214  
FILING DATE: 10-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1110-139P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-339-214-8

Query Match 100.0%; Score 50; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.059;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKAELEST 11  
DB 109 KEKAELEST 119

RESULT 9  
US-08-339-214-30  
Sequence 30, Application US/08339214  
Patent No. 6348334  
GENERAL INFORMATION:

APPLICANT: Nagata, Shigikazu  
APPLICANT: Suda, Takashi  
APPLICANT: Takahashi, Tomoniro  
APPLICANT: Nakamura, No. 634833410  
TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/339,214  
FILING DATE: 10-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1110-139P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-339-214-30

Query Match 100.0%; Score 50; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.059;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKAELEST 11  
DB 109 KEKAELEST 119

RESULT 10  
US-09-589-287B-6  
Sequence 6, Application US/09589287B  
Patent No. 6403770  
GENERAL INFORMATION:  
APPLICANT: Yu et al.  
TITLE OF INVENTION: Antibodies to Neutrokin- $\alpha$   
FILE REFERENCE: PF343P3C1  
CURRENT APPLICATION NUMBER: US/09/589,287B  
CURRENT FILING DATE: 2000-06-08  
Prior application data removed - check PALM or file wrapper  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 6  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Description of Combined DNA/RNA Molecule: n equals  
US-09-589-287B-6

Query Match 100.0%; Score 50; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.059;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKAELEST 11  
DB 109 KEKAELEST 119

RESULT 11  
US-09-157-864-10  
Sequence 10, Application US/09157864  
Patent No. 6440694  
GENERAL INFORMATION:  
APPLICANT: Bienkowski, Michael J  
APPLICANT: Mills, Cynthia J  
APPLICANT: Jones, David A  
TITLE OF INVENTION: TNF-Related Death Ligand  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pharmacia & Upjohn, Intellectual Property  
ADDRESSEE: Legal Services  
STREET: 301 Henrietta Street  
CITY: Kalamazoo  
STATE: MI  
COUNTRY: USA  
ZIP: 49001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 Diskette  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/157,864  
FILING DATE:

```

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kerber, Lori L.
; REGISTRATION NUMBER: 41,113
; REFERENCE/DOCKET NUMBER: 6111.N CNI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 616/833-0974
; TELEFAX: 616/833-8897
; TELEX: 224401
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-157-864-10

Query Match          100.0%; Score 50; DB 4; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KELAELRESTS 11
        |||||
Db      109 KELAELRESTS 119

RESULT 12
US-09-006-755B-1
; Sequence 1, Application US/09006755B
; Patent No. 6451759
; GENERAL INFORMATION:
; APPLICANT: Kang, Sang-MO
; APPLICANT: Braet, Andries E
; APPLICANT: Baekeskov, Steinunn
; APPLICANT: Stock, Peter G.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A No. 6451759cleavable Fas ligand
; FILE REFERENCE: 18062K-000500US
; CURRENT APPLICATION NUMBER: US/09/006,755B
; CURRENT FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (103)..(281)
; OTHER INFORMATION: extracellular domain
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (119)..(154)
; OTHER INFORMATION: metalloprotease recognition region
; FEATURE:
; OTHER INFORMATION: human wild-type Fas ligand (FasL)
; US-09-006-755B-1

Query Match          100.0%; Score 50; DB 4; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KELAELRESTS 11
        |||||
Db      109 KELAELRESTS 119

RESULT 13
PCT-US95-00362-2
; Sequence 2, Application PC/TUS9500362
; GENERAL INFORMATION:
; APPLICANT: IMMUNEX CORPORATION
```

```

; TITLE OF INVENTION: Ligand That Binds Fas Antigen
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple 7.1
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/00362
; FILING DATE: 06-JAN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/179,138
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,559
; FILING DATE: 01-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Anderson, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2805-MO
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-00362-2

Query Match          100.0%; Score 50; DB 5; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KELAELRESTS 11
        |||||
Db      109 KELAELRESTS 119

RESULT 14
US-08-815-190A-16
; Sequence 16, Application US/08815190A
; Patent No. 6046310
; GENERAL INFORMATION:
; APPLICANT: Queen, Cary L.
; APPLICANT: Schneider, William P.
; APPLICANT: Vasquez, Maximiliano
; TITLE OF INVENTION: Fas ligand Fusion Proteins and Their
; TITLE OF INVENTION: Fas ligand Fusion Proteins and Their
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/815,190A
; FILING DATE: 11-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
```

Search completed: March 13, 2003, 18:53:50  
job time : 10.8 secs

APPLICATION NUMBER: US 08/614,584  
FILING DATE: 13-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 011823-0067100S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 287 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-815-190A-16

Query Match 100.0%; Score 50; DB 3; Length 287;  
Best Local Similarity 100.0%; Pred. No. 0.06;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
DB 115 KELAELRESTS 125

RESULT 15  
US-08-751-512-8  
Sequence 8, Application US/08751512  
Patent No. 6001962  
GENERAL INFORMATION:  
APPLICANT: Ramer, J. Kevin  
APPLICANT: Williams, Lewis T.  
TITLE OF INVENTION: Modified FAS ligands  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/751,512  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Matthew B.  
REGISTRATION NUMBER: 39,787  
REFERENCE/DOCKET NUMBER: 02307K-07100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-751-512-8

Query Match 100.0%; Score 50; DB 3; Length 376;  
Best Local Similarity 100.0%; Pred. No. 0.079;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
DB 204 KELAELRESTS 214



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:40:28 ; Search time 5.8 seconds

(without alignments)  
78.662 Million cell updates/sec

Title: US-10-050-200-3

Sequence: 1 KRLAELRESTS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSPROT\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	280	TNF6_CERTO	Q9bdn1 cercocebus
2	50	100.0	280	TNF6_MACMU	Q9myv1 macaca mula
3	50	100.0	281	TNF6_HUMAN	P48073 homo sapien
4	41	82.0	278	TNF6_RAT	P36940 rattus norv
5	41	82.0	279	TNF6_MOUSE	P41047 mus musculi
6	41	82.0	282	TNF6_PIG	Q9bea8 sus scrofa
7	38	76.0	382	SUC1_ARCFU	O28732 archaeoglob
8	37	74.0	408	VLEN_NPVAC	P24647 autographa
9	35	70.0	370	T2M2_METJA	O58844 methanococ
10	35	70.0	1333	YN95_YEAST	P53786 saccharomyc
11	34	68.0	673	Y552_HUMAN	O60299 homo sapien
12	34	68.0	1041	SY1_MYCTU	O10765 mycobacteri
13	34	68.0	1585	YOBO_BACSU	P45931 bacillus su
14	33	66.0	138	HEX9_ADE07	P03283 human adeno
15	33	66.0	293	MT16_EMENTI	P56859 emeticella
16	33	66.0	295	SAPC_HAEIN	P45287 haemophilus
17	33	66.0	387	SUCC_CWJME	Q9phv1 campylobact
18	33	66.0	405	VIE2_NPVOL	Q9phj1 oryza pseu
19	33	66.0	464	YRG5_CAEEL	O09406 caenorhabdi
20	33	66.0	473	YVCE_BACSU	P40767 bacillus su
21	33	66.0	576	Y065_CHLPN	Q92966 chlamydia p
22	33	66.0	607	DNK_MTXXA	P95334 myxococcus
23	33	66.0	1162	CAR8_AGR75	Q8ude9 agrobacteri
24	32	64.0	72	RL29_CHIMU	Q9phm2 chlamydia m
25	32	64.0	72	RL29_CHLTR	P28538 chlamydia t
26	32	64.0	149	CYNS_AQUAE	O66587 aquifex aeo
27	32	64.0	174	GRPE_METTH	O27350 methanobact
28	32	64.0	180	YMW6_BACSU	P39157 bacillus su
29	32	64.0	235	CAMT_POPKI	P93711 populus kit
30	32	64.0	250	ODFP_HOVIN	Q14990 homo sapien
31	32	64.0	262	ODFP_BOVIN	O29438 bos taurus
32	32	64.0	262	ODFP_PIG	O29077 sus scrofa
33	32	64.0	323	TTD2_HUMAN	P09758 homo sapien

## ALIGNMENTS

RESULT 1	TNF6_CERTO	STANDARD;	PRT;	280 AA.
AC	Q9BDN1:			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)			
DE	(CD95L protein).			
GN	TNFSF6 OR FASL OR CD95L.			
OS	Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;			
OC	Cercopitheciinae; Cercopithecus.			
OX	NCBI_TaxID=9531;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymphocytes;			
RX	MEDLINE-21383618: Pubmed-11491535;			
RA	Vallinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,			
RA	Weiss W.R., Ansari A.A.;			
RT	'Cloning, sequencing, and homology analysis of nonhuman primate			
RT	Fas/FasL ligand and co-stimulatory molecules.';			
RL	Immunogenetics 53:315-328(2001).			
-	FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that			
-	transduces the apoptotic signal into cells. May be involved in			
-	cytotoxic T cell mediated apoptosis and in T cell development.			
-	TNFRSF6/FAS-mediated apoptosis may have a role in the induction of			
-	peripheral tolerance, in the antigen-stimulated suicide of mature			
-	T cells, or both. Binding to the decoy receptor TNFRSF6/DCR3			
-	modulates its effects (By similarity).			
-	SUBUNIT: Homotrimer (Probable).			
-	SUBCELLULAR LOCATION: Type II membrane protein and secreted (By			
-	similarity).			
-	PTM: The soluble form derives from the membrane form by			
-	proteolytic processing (By similarity).			
-	SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.			
-	-----			
-	This SWISS-PROT entry is copyright. It is produced through a collaboration			
-	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
-	the European Bioinformatics Institute. There are no restrictions on its			
-	use by non-profit institutions as long as its content is in no way			
-	modified and this statement is not removed. Usage by and for commercial			
-	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
-	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
-	-----			
-	EMBL: AF344847; AKR37606.1; -			
-	HSSP: P01375; ATSV.			
-	InterPro: IPR003636; TNF_abc.			
-	InterPro: IPR000478; TNF_family.			
-	Pfam: PF00229; TNF_1.			
-	PRINTS: PR01234; TNECROSISFCT.			
-	PRODOM: PD002012; TNF_abc; 1.			
-	SMART: SM00207; TNF_1.			
-	SMART: PS00251; TNF_1; 1.			
-	PROSITE: PS50049; TNF_2; 1.			

34	32	64.0	407	1	RFL_ARCFU	O29048 archaeoglob
35	32	64.0	413	1	NUOD_RHOCA	O07310 rhodobacter
36	32	64.0	502	1	GLPK_PASMU	P57944 pasteurella
37	32	64.0	509	1	YRM4_CAEEL	Q10051 caenorhabdi
38	32	64.0	623	1	HPC2_YEAST	Q01448 saccharomyc
39	32	64.0	688	1	SYGB_ECOLI	P00961 escherichia
40	32	64.0	715	1	HEL5_PYPAB	Q9v089 pyrococcus
41	32	64.0	1501	1	SNO2_YEAST	P32568 saccharomyc
42	32	64.0	2347	1	KROS_HUMAN	P08922 homo sapien
43	31	62.0	57	1	GNS_SALTY	O8xge3 salmonella
44	31	62.0	120	1	YQXJ_BACSU	P24809 bacillus su
45	31	62.0	126	1	PFDB_SULSO	Q9uxb8 sulfolobus

```

KM Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 280
    TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
    MEMBER 6, MEMBRANE FORM.
FT CHAIN 129 280
    TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
    MEMBER 6, SOLUBLE FORM (BY SIMILARITY).
FT DOMAIN 1 80
    CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 101
    SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
    (POTENTIAL).
FT DOMAIN 102 280
    EXTRACELLULAR (POTENTIAL).
FT DOMAIN 4 69
    PRO-RICH.
FT DOMAIN 128 129
    POLY-PRO.
FT SITE 128 129
    CLEAVAGE (BY SIMILARITY).
FT DISULFID 201 232
    POTENTIAL.
FT CARBOHYD 183 183
    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 249 249
    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259
    N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 280 AA; 31407 MW; 729EA6067B7D398 CRC64;

Query Match 100.0%; Score 50; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11
Db 108 KELAIRESTS 118

RESULT 2
TNF6_MACMU STANDARD: PRT: 280 AA.
ID. TNF6_MACMU
AC Q9MYL6; Q9BDM5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)
DE (CD95L protein)
GN TNFRSF6 OR FASL OR CD95L.
OS Macaca mulatta (Rhesus macaque),
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Macaca nemestrina (Pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Cercopithecoidea;
OC Cercopithecoidea; Macaca
OX NCBI_TaxID=9344, 9341, 9345;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M. mulatta; TISSUE=Lymphocytes;
RX MEDLINE=21383618; PubMed=11491535;
RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,
RA Weiss W.R., Ansari A.A.;
RA "Cloning, sequencing, and homology analysis of nonhuman primate
RA Fas/Fas-ligand and co-stimulatory molecules.";
RT Immunogenetics 53:315-328(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M. mulatta, M. fascicularis, and M. nemestrina;
RA Kiril Y., Inoue T., Yoshino K.;
RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that
CC transduces the apoptotic signal into cells. May be involved in
CC cytotoxic T cell mediated apoptosis and in T cell development.
CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of
CC peripheral tolerance, in the antigen-stimulated suicide of mature
CC T cells, or both. Binding to the decoy receptor TNFRSF6B/Dec3
CC modulates its effects (By similarity).
CC -1- SUBUNIT: Homotrimer (Potential).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -1- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF344856; AAK37539.1; -
DR EMBL: AB035138; BAA90294.1; -
DR EMBL: AB035139; BAA90295.1; -
DR EMBL: AB035140; BAA90296.1; -
DR HSP: P01375; 4TSY.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR00478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR PRINTS: PR01234; TNECROSISFCT.
DR PRODOM: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KM Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 280
    TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
    MEMBER 6, MEMBRANE FORM.
FT CHAIN 129 280
    TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
    MEMBER 6, SOLUBLE FORM (BY SIMILARITY).
FT TRANSMEM 81 101
    SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
    (POTENTIAL).
FT DOMAIN 102 280
    EXTRACELLULAR (POTENTIAL).
FT DOMAIN 4 69
    PRO-RICH.
FT DOMAIN 128 129
    POLY-PRO.
FT SITE 128 129
    CLEAVAGE (BY SIMILARITY).
FT DISULFID 201 232
    POTENTIAL.
FT CARBOHYD 183 183
    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 249 249
    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259
    N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 60 60
    S -> P (IN REF. 1).
SQ SEQUENCE 280 AA; 31367 MW; F0B284D61A132EB4 CRC64;

Query Match 100.0%; Score 50; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11
Db 108 KELAIRESTS 118

RESULT 3
TNF6_HUMAN STANDARD: PRT: 281 AA.
ID. TNF6_HUMAN
AC P48023; Q9BZP9;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)
DE (Apoptosis antigen ligand) (APTL) (CD178 antigen).
GN TNFRSF6 OR FASL OR APTL1G1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC MEDLINE=95105731; PubMed=7528780;
RA Alderson M.;
RA "Fas ligand mediates activation-induced cell death in human T
RA lymphocytes.";
RL J. Exp. Med. 181:71-77(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC MEDLINE=95127560; PubMed=7826947;
RX Takahashi T., Tanaka M., Inazawa J., Abe T., Suda T., Nagata S.;

```

RT "Human Fas ligand: gene structure, chromosomal location and species  
 RT specificity.";  
 RL Int. Immunol. 6:1567-1574(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Schetzle C.E., Pöhlmann R., Philippson P., Eibel H.;  
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA MEDLINE=95071350; PubMed=7980502;  
 RA Mita E., Hayashi N., Iio S., Takehara T., Hijioka T., Kasahara A.,  
 RA Fusamoto H., Kamada T.;  
 RT "Role of Fas ligand in apoptosis induced by hepatitis C virus  
 RT infection";  
 RL Biochem. Biophys. Res. Commun. 204:468-474(1994).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Leukocyte;  
 RA Zeytun A., Nagarkatti M., Nagarkatti P.S.;  
 RT "Isolation and characterization of a new naturally occurring variant of  
 RT human Fas ligand that is expressed only in membrane bound form.";  
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Wilkinson J.;  
 RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Blood;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.  
 RN [8]  
 RP SEQUENCE OF 1-10 FROM N.A.  
 RC TISSUE=Blood;  
 RA Matsumura M., Nakanishi Y., Ohba Y.;  
 RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.  
 RN [9]  
 RP CHARACTERIZATION, AND MUTAGENESIS OF PRO-206, TYR-218 AND PHE-275.  
 RA Schneider P., Bodmer J.-L., Holler N., Mattmann C., Scuderi P.,  
 RA Terziklik A., Fetsch M.C., Tschopp J.;  
 RT "Characterization of Fas (Apo-1, CD95)-Fas ligand interaction.";  
 RL J. Biol. Chem. 272:18827-18833(1997).  
 RN [10]  
 RP PROCESSING.  
 RA Tanaka M., Itai T., Adachi M., Nagata S.;  
 RT "Downregulation of Fas ligand by shedding.";  
 RL Nat. Med. 4:31-36(1998).  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF6/Fas, a receptor that  
 CC transduces the apoptotic signal into cells. May be involved in  
 CC cytotoxic T cell mediated apoptosis and in T cell development.  
 CC TNFRSF6/Fas-mediated apoptosis may have a role in the induction of  
 CC peripheral tolerance. In the antigen-stimulated suicide of mature  
 CC T cells, or both. Binding to the decoy receptor TNFRSF6B/Dcr3  
 CC modulates its effects.  
 CC -1- SUBUNIT: HOMODIMER (PROBABLY).  
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED  
 CC INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FROM THE CELL  
 CC SURFACE.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: 1 (shown here) and 2;  
 CC are produced by alternative splicing.  
 CC -1- PTM: N-glycosylated.  
 CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing.  
 CC -1- DISEASE: Defects in TNFRSF6 are a cause of autoimmune  
 CC lymphoproliferative syndrome (ALPS), also known as Canale-Smith  
 CC syndrome (CSS), a childhood syndrome involving hemolytic anemia  
 CC and thrombocytopenia with massive lymphadenopathy and  
 CC splenomegaly.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- DATABASE: NAME=PROW; NOTE=PROW 2:59-69(2001);  
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/3338769674\_g.htm".

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X89102; CAA61474.1; -  
 CC EMBL: U08137; AAC50071.1; -  
 CC EMBL: U11821; AAC50124.1; -  
 CC EMBL: D38122; BAA07320.1; -  
 CC EMBL: AF288573; AAG60017.1; -  
 CC EMBL: Z96050; CAB09424.1; -  
 CC EMBL: BC017502; AAH1502.1; -  
 CC EMBL: AB013303; BAA32542.1; -  
 CC HSSP: P01375; ITNF.  
 CC Genev: HGNC:11936; TNFRSF6.  
 CC MIM: 134638; -  
 CC MIM: 601859; -  
 CC InterPro: IPR003636; TNF\_abc.  
 CC InterPro: IPR000478; TNF\_family.  
 CC Pfam: PF00229; TNF; 1.  
 CC ProDom: PD002012; TNF\_abc; 1.  
 CC SMART: SM00207; TNF; 1.  
 CC PROSITE: PS00251; TNF\_1; 1.  
 CC PROSITE: PS50049; TNF\_2; 1.  
 CC KMW Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor;  
 CC Alternative splicing; Antigen.  
 CC CHAIN 1 281  
 CC FT CHAIN 130 281  
 CC FT DOMAIN 1 80  
 CC FT TRANSMEM 81 102  
 CC FT DOMAIN 103 281  
 CC FT DOMAIN 4 70  
 CC FT SITE 129 130  
 CC FT DISULFID 202 233  
 CC FT CARBOHYD 184 184  
 CC FT CARBOHYD 250 250  
 CC FT CARBOHYD 260 260  
 CC FT VARSPPLIC 117 127  
 CC FT VARSPPLIC 128 281  
 CC FT MUTAGEN 206 206  
 CC FT MUTAGEN 218 218  
 CC FT MUTAGEN 275 275  
 CC FT SEQUENCE 281 AA; 31485 MW; A8A6EB558246EBB CRC64;  
 CC Query Match 100.0%; Score 50; DB 1; Length 281;  
 CC Best Local Similarity 100.0%; Pred. No. 0.032;  
 CC Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC DB 109 KEIARESTS 119  
 CC QY 1 KEIARESTS 11  
 CC ID 109 KEIARESTS 119  
 CC RESULT 4  
 CC TNF6\_RAT  
 CC AC P36940; STANDARD; PRT; 278 AA.  
 CC DT 01-JUN-1994 (Rel. 29, Created)  
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 6 (Fas antigen  
ligand).  
GN TNFSF6 OR APTL1G1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A. PubMed=7505205;  
RX MEDLINE=94084792; Nagata S.;  
RA Suda T., Takehashi T., Golstein P., Nagata S.;  
RT "Molecular cloning and expression of the Fas ligand, a novel member  
of the tumor necrosis factor family.";  
RL Cell 75:1169-1178(1993).  
CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that  
transduces the apoptotic signal into cells. May be involved in  
cytotoxic T cell mediated apoptosis and in T cell development.  
CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of  
peripheral tolerance, in the antigen-stimulated suicide of mature  
T cells, or both. Binding to the decoy receptor TNFRSF6b/DCR3  
modulates its effects (By similarity).  
CC -1- SUBUNIT: HOMOTRIMER (PROBABLE).  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By  
similarity).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ACTIVATED SPLENOCYTES AND  
THYMOCYTES. MODERATE OR WEAK EXPRESSION FOUND IN SMALL INTESTINES,  
KIDNEY AND LUNG.  
CC -1- INDUCTION: BY PMA/TIONOMYCIN AND CONCAVALIN/INTERLEUKIN-2.  
CC -1- PTM: The soluble form derives from the membrane form by  
proteolytic processing (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib.ch](mailto:license@sib.ch)).  
CC -----  
DR EMBL: U03470; AAC52129.1; -.  
DR HSSP: P01375; 4TSV.  
DR InterPro: IPR003636; TNF\_abc.  
DR InterPro: IPR000478; TNF\_family.  
DR Pfam: PF00229; TNF; 1.  
DR ProDom: PD002012; TNF\_abc; 1.  
DR SMART: SM00207; TNF; 1.  
DR PROSITE: PS00251; TNF\_1; 1.  
DR PROSITE: PS50049; TNF\_2; 1.  
KW Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
FT CHAIN 1 278 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
FT CHAIN 127 278 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
FT DOMAIN 1 77 MEMBER 6, SOLUBLE FORM (BY SIMILARITY).  
FT TRANSMEM 78 99 CYTOPLASMIC (POTENTIAL).  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT SIGNAL-ANCHOR (POTENTIAL).  
FT DOMAIN 100 278 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 4 69 PRO-RICH.  
FT DOMAIN 45 58 POLY-PRO.  
FT SITE 126 127 CLEAVAGE (BY SIMILARITY).  
FT DISULFID 199 230 POTENTIAL.  
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 278 AA; 31140 MM; 28981B62CEAC6 CRC64;  
Query Match Score 41; DB 1; Length 278;  
Best Local Similarity 81.8%; Pred. No. 1.5;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
OY 1 KELAELRESTS 11  
||||||| 1:

DB 106 KELAELREFTN 116  
RESULT 5  
TNF6\_MOUSE STANDARD: PRT; 279 AA.  
AC P41047; Q61217; Q9RIE2;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tumor necrosis factor ligand superfamily member 6 (Fas antigen  
ligand).  
GN TNFSF6 OR APTL1G1 OR GLD.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM FASL).  
RX MEDLINE=94185175; PubMed=7511063;  
RA Takehashi T., Tanaka M., Brannan C.I., Jenkins N.A., Copeland N.G.,  
RA Suda T., Nagata S.;  
RT "Generalized lymphoproliferative disease in mice, caused by a point  
mutation in the Fas ligand.";  
RL Cell 76:969-976(1994).  
[2]  
RP SEQUENCE FROM N.A. (ISOFORM FASL), AND 3D-STRUCTURE MODELING.  
RC STRAIN=C57BL/6;  
RX MEDLINE=95388076; PubMed=7544870;  
RA Peitsch M.J., Tschopp J.J.;  
RT "Comparative molecular modelling of the Fas-ligand and other members  
of the TNF family.";  
RL Mol. Immunol. 32:761-772(1995).  
[3]  
RP SEQUENCE FROM N.A. (ISOFORM FASL).  
RX MEDLINE=95196085; PubMed=789405;  
RA Lynch D.H., Watson M.L., Alderson M.R., Baum P.R., Miller R.E.,  
RA Tough T., Gibson M., Davis-Smith T., Smith C.A., Hunter K.;  
RT "The mouse Fas-ligand gene is mutated in gld mice and is part of a  
TNF family gene cluster.";  
RL Immunity 1:131-136(1994).  
[4]  
RP SEQUENCE FROM N.A. (ISOFORM FASL).  
RC STRAIN=BALB/c;  
RA Fenner M.H., Shioda T., Isselbacher K.J.;  
RT "Mus musculus Balb/c Fas ligand differs from 129/SV Fas ligand in  
two amino acids.";  
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.  
[5]  
RP SEQUENCE FROM N.A. (ISOFORM FASLS).  
RC STRAIN=C3H; TISSUE=Spleen;  
RX MEDLINE=20021694; PubMed=10552956;  
RA Ayroldi E., D'Adamo F., Zollo O., Agostini M., Moraca R.,  
RA Camarillo L., Migliorati G., Delino D.V., Riccardi C.;  
RT "Cloning and expression of a short Fas ligand: A new alternatively  
spliced product of the mouse Fas ligand gene.";  
RL Blood 94:3456-3467(1999).  
[6]  
RP CHARACTERIZATION OF VARIANT GLD.  
RX MEDLINE=96091792; PubMed=7495745;  
RA Hahne M., Peitsch M.C., Imler M., Schroeder M., Lowin B.,  
RA Rousseau M., Bron C., Renno T., French L., Tschopp J.;  
RT "Characterization of the non-functional Fas ligand of gld mice.";  
RL Int. Immunol. 7:1381-1386(1995).  
[7]  
RP VARIANTS ALA-184 AND GLY-218.  
RC STRAIN=Various;  
RX PubMed=9108079;  
RA Kayagaki N., Yamaguchi N., Nagao F., Matsuo S., Maeda H., Okumura K.,  
RA Tagita H.;  
RT "Polymorphism of murine Fas ligand that affects the biological  
activity.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:3914-3919(1997).

```

CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that
CC transduces the apoptotic signal into cells. May be involved in
CC cytotoxic T cell mediated apoptosis and in T cell development.
CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of
CC peripheral tolerance. In the antigen-stimulated suicide of mature
CC T cells, or both. Binding to the decoy receptor TNFRSF6B/DCR3
CC modulates its effects (By similarity).
CC -1- SUBUNIT: HOMODIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (isoform FASL);
CC secreted (isoforms FASL and FASL5).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: FASL (SHOWN HERE) AND FASL5; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. ISOFORM FASL MEDIATES APOPTOSIS
CC WHILE ISOFORM FASL5 PREVENTS APOPTOSIS INDUCED BY FAS/FASL
CC INTERACTION.
CC -1- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -1- DISEASE: A DEFICIENCY IN THIS PROTEIN IS THE CAUSE OF GENERALIZED
CC LYMPHOPROLIFERATION DISEASE (GID), AN AUTOSOMAL RECESSIVE DISEASE
CC RESPONSIBLE FOR LYMPHADENOPATHY AND AUTOMATOID PRODUCTION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: 006948; AAA17800.1; -
CC EMBL: U10984; AAA19778.1; -
CC EMBL: S76752; AAB33780.1; -
CC EMBL: U58995; AAB02915.1; -
CC EMBL: AF119335; AAB52106.1; -
CC HSSP: P01375; 4TSV.
CC MGD: MGI:99255; Tnf:6.
CC InterPro: IPR003636; TNF_abc.
CC InterPro: IPR000478; TNF_family.
CC Pfam: PF00229; TNF_1.
CC ProDom: PD002012; TNF_abc; 1.
CC SMART: SM00207; TNF; 1.
CC PROSITE: PS00251; TNF_1; 1.
CC PROSITE: PS50049; TNF_2; 1.
CC KEGG: Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor;
CC Disease mutation; Polymorphism;
CC CHAIN 1 279
CC CHAIN 128 279
CC CHAIN 1 79
CC DOMAIN 1 78
CC TRANSMEM 4 69
CC DOMAIN 101 279
CC DOMAIN 4 69
CC SITE 127 128
CC SITE 127 128
CC DISULFID 200 231
CC CARBOHYD 117 117
CC CARBOHYD 182 182
CC CARBOHYD 248 248
CC CARBOHYD 258 258
CC VARSPLIC 1 210
CC VARIANT 184 184
CC VARIANT 218 218
CC VARIANT 273 273
CC VARIANTE 273 273
CC SEQUENCE 279 AA; 31442 MW; 37972E2728B0A1CA CRC64;
Query Match 82.0%; Score 41; DB 1; Length 279;
Best Local Similarity 81.8%; Pred. No. 1.5;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 KETAPLESTS 11
Db 107 KETAPLESTFN 117
RESULT 6
TNF6_PIG STANDARD: PRT; 282 AA.
AC Q9BEA8; Q95N10; Q95M04;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 6 (Fas antigen
DE ligand).
DE TNFRSF6 OR FASL.
GN Sus scrofa (Pig).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI:TaxID=9823;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-2132533; PubMed-11429161;
RA Muneta Y., Shimoi Y., Inumaru S., Mori Y.;
RT "Molecular cloning, characterization, and expression of porcine Fas
RT ligand (CD95 ligand).";
RL J. Interferon Cytokine Res. 21:305-312(2001).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-Guanxi lama miniature pig;
RA Zhu N., Young Y.;
RT "Molecular cloning and characterization of porcine Fas ligand cDNA.";
RN Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC TISSUE-Lymphoid;
RA Tsuyuki S., Kono M., Bloom E.T.;
RT "Cloning and potential utility of porcine Fas ligand: overexpression
RT in porcine cells protects them from attack by human cytolytic cells.";
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN-Landrace x Large Yorkshire white; TISSUE-Thymocytes;
RX PubMed-11792426;
RA Motegi-Ishiyama Y., Nakajima Y., Hoka S., Takagaki Y.;
RT "Porcine Fas-ligand gene: genomic sequence analysis and comparison
RT with human gene.";
RL Mol. Immunol. 38:581-586(2002).
CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that
CC transduces the apoptotic signal into cells. May be involved in
CC cytotoxic T cell mediated apoptosis and in T cell development.
CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of
CC peripheral tolerance. In the antigen-stimulated suicide of mature
CC T cells, or both. Binding to the decoy receptor TNFRSF6B/DCR3
CC modulates its effects (By similarity).
CC -1- SUBUNIT: Homodimer (Probable).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -1- INDUCTION: By IL-18.
CC -1- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AB027297; BAB40919.1; -
CC EMBL: AV033634; AAK56449.1; -

```

DR EMBL: AF397407; AAK84408.1; -;  
 DR EMBL: AB069764; BAB64291.1; -;  
 DR HSSP: P01375; 4TSV.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR004478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR PRINTS: PR01234; TNECROSISFCT.  
 DR PRODOM: PDO02012; TNF\_abc; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KM Cytokine: Apoptosis; TNF\_2; 1.  
 FT CHAIN 1  
 FT CHAIN 1 282  
 FT CHAIN 131 282  
 FT CHAIN 1 82  
 FT TRANSMEM 83 103  
 FT DOMAIN 104 282  
 FT DOMAIN 4 70  
 FT DOMAIN 45 56  
 FT SITE 130 131  
 FT DISULFID 203 234  
 FT CARBOHYD 185 185  
 FT CARBOHYD 251 251  
 FT CARBOHYD 261 261  
 FT CONFLICT 5 5  
 FT CONFLICT 57 57  
 SO: SEQUENCE 282 AA; 31756 MW; 6743DA1145671FB CRC64;

Query Match  
 Best Local Similarity 82.0%; Score 41; DB 1; Length 282;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KELAIRESTS 11  
 Db 110 KELAIRESTS 120

RESULT 7  
 SUC1\_ARCFU STANDARD; PRT; 382 AA.  
 AC 028732;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Succinyl-CoA synthetase beta chain 1 (EC 6.2.1.5) (SCS-beta 1).  
 GN SUC1 OR AF1540.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 OX NCBI\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475;  
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwyn M., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kertevage A.R., Graham D.E., Kyriades N.C.,  
 RA Pleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kitznes E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Relch C.I., McNeill L.K., Badger J.H., Glodok A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,  
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,  
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason J.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 reducing archaeon Archaeoglobus fulgidus.";  
 RL Nature 390:364-370(1997).  
 CC -1- CATALYTIC ACTIVITY: ATP + succinate + CoA = ADP + succinyl-CoA +  
 phosphate.

CC -1- PATHWAY: Tricarboxylic acid cycle.  
 CC -1- SUBUNIT: Composed of an alpha chain and a beta chain (By  
 CC similarity).  
 CC -1- SIMILARITY: BELONGS TO THE SUCCINYL-CoA SYNTHETASE BETA SUBUNIT  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AE000996; AAB89706.1; ALU\_INIT.  
 DR HSSP: P07460; 2SCU.  
 DR TIGR: AF1540; -;  
 DR InterPro: IPR003135; ATP-grasp.  
 DR InterPro: IPR000303; CoA\_ligase.  
 DR Pfam: PF00549; ligase-CoA; 1.  
 DR Pfam: PF02222; ATP-grasp; 1.  
 DR TIGRfams: TIGR01016; succCoBeta; 1.  
 DR PROSITE: PS01217; SUCCINYL-CoA\_Lig\_3; 1.  
 KM Ligase; Tricarboxylic acid cycle; Complete proteome.  
 SO SEQUENCE 382 AA; 42204 MW; 418ACBA41E64DA9F CRC64;

Query Match  
 Best Local Similarity 76.0%; Score 38; DB 1; Length 382;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ELAELREST 10  
 Db 219 ELAELREST 227

RESULT 8  
 VIEN\_NPVAC STANDARD; PRT; 408 AA.  
 AC P24647;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Immediate-early regulatory protein IE-N.  
 GN IE-N.  
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 OC Nucleopolyhedrovirus.  
 OX NCBI\_TaxID=46015;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=E2;  
 RX MEDLINE=91220660; PubMed=2024466;  
 RA Carson D.D., Summers M.D., Guarino L.A.;  
 RT Molecular analysis of a baculovirus regulatory gene.";  
 RL Virology 182:279-286(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C6;  
 RX MEDLINE=94303173; PubMed=8030224;  
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;  
 RT "The complete DNA sequence of Autographa californica nuclear  
 RT polyhedrosis virus.";  
 RL Virology 202:586-605(1994).  
 CC -1- FUNCTION: IE-N AUTOREGULATES ITS OWN EXPRESSION AND STIMULATES  
 CC BOTH IE-1 AND IE-0 IN TRANSIENT ASSAYS. THE STIMULATION OF IE-1  
 CC EXPRESSION MAY ACCOUNT FOR THE AUGMENTING ACTIVITY OF IE-N IN THE  
 CC IE-1 MEDIATED TRANS-ACTIVATION OF THE 39K PROMOTER.  
 CC -1- SIMILARITY: TO OPMNPV IMMEDIATE-EARLY PROTEIN IE-2.  
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59422; AAA46701.1; -.
DR EMBL; L23858; AAA66781.1; -.
DR PIR; A39150; KMWYIA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; Ring; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Early protein; Transcription regulation; Trans-acting factor;
KM Activator; Repeat; Zinc-finger; DNA-binding.
FT DOMAIN 9
FT DOMAIN 34 49
FT DOMAIN 51 58
FT DOMAIN 161 189
FT DOMAIN 188 196
FT DOMAIN 198 204
FT ZN_FING 207 255
FT CONFLICT 138 138
SQ SEQUENCE 408 AA; 47007 MW; DDB6F486ZCCD8553 CRC64;
Query Match 74.0%; Score 37; DB 1; Length 408;
Best Local Similarity 72.7%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 KELAELRESTS 11
Db 311 KELSRLAKTS 321
RESULT 9
T2M2_METUA
ID T2M2_METUA STANDARD; PRT; 370 AA.
AC Q58844;
DT 16-OCT-2001 (Rel. 40, Created)
DP 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Type II restriction enzyme MjaII (EC 3.1.21.4) (Endonuclease MjaII)
DE (R MjaII)
GN MAIIR OR MJ1449.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.L.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RT Science 273:1058-1073(1996).
[2]
RN CHARACTERIZATION.
RP Zetler B., Myers P.A., Escalante H., Roberts R.J.;
RL Unpublished Observations (XXX-1997).
CC -1- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE GCNCC.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give
CC specific double-stranded fragments with terminal 5'-phosphates.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```

CC	the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL: U67585; AAB99461.1; -.
DR	REBASE: 1222; Mjail.
DR	TIGR: Mj1449; -.
KW	Hydrolase; Endonuclease; Nuclease; Restriction system; Complete proteome.
SO	SEQUENCE 370 AA; 43911 MW; 8B4CF6208F914B33 CRC64;
QY	1 KEIARESTS 11
Db	126 KENKLEAETS 136
RESULT 10	
YN99_YEAST	
ID	YN99_YEAST STANDARD; PRT: 1333 AA.
AC	P53756.
DT	01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	Probable ATP-dependent transporter YNR070W.
GN	YNR070W ORF N3568.
OS	Saccharomyces cerevisiae (Baker's yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX	NCHI_TaxID=4932;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Andre B., Iragui Houssaini I., Urrestiarazu L.A., Vissers S.;
RL	Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE OF 1-1054 FROM N.A.
RA	Duesterhoeft A., Floeth M., Filtz C., Heuss-Netzel D.,
RL	Hilbert H., Moestl D.;
RL	Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. PDR5 SUBFAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL: Z71685; CAA96352.1; -.
DR	EMBL: Z71686; CAA96354.1; -.
DR	SGD: S0005353; YNR070W.
DR	InterPro: IPR003593; AAA_Arpase.
DR	InterPro: IPR003439; ABC_transportr.
DR	Pfam: PF00005; ABC_tran; 2.
DR	ProDom: PD000006; ABC_transportr; 2.
DR	SMART: SM00382; AAA; 1.
DR	PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW	Hypothetical protein; ATP-binding; Transmembrane; Glycoprotein;
KW	Transport.
KW	TRANSMEM 13
FT	TRANSMEM 392 412 POTENTIAL.
FT	TRANSMEM 425 445 POTENTIAL.
FT	TRANSMEM 474 494 POTENTIAL.
FT	TRANSMEM 499 519 POTENTIAL.
FT	TRANSMEM 534 554 POTENTIAL.
FT	TRANSMEM 642 662 POTENTIAL.

```

FT TRANSMEM 1071 1091 POTENTIAL.
FT TRANSMEM 1092 1112 POTENTIAL.
FT TRANSMEM 1150 1170 POTENTIAL.
FT TRANSMEM 1178 1198 POTENTIAL.
FT TRANSMEM 1210 1230 POTENTIAL.
FT TRANSMEM 1235 1255 POTENTIAL.
FT NP_BIND 765 772 ATP (POTENTIAL).
FT CAROHD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 697 697 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 958 958 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 1320 1320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1333 AA; 149749 MW; 6184758E0245CB70 CRC64;

Query Match
Best Local Similarity 70.0%; Score 35; DB 1; Length 1333;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RELATERESTS 11
DB 1122 KELVEPRESK 1132

RESULT 11
Y552_HUMAN STANDARD; PRT; 673 AA.
AC 060299; 15-JUL-1999 (Rel. 38, Created)
BT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA0552.
GV KIAA0552.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT *Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.*;
RL DNA Res. 5:31-39(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clisg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Garmham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehtvaaho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachle L.J., McElay K., McKerrry A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

```

```

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT *The DNA sequence and comparative analysis of human chromosome 20.*;
RL Nature 414:865-871(2001).
CC -I- SIMILARITY: TO HUMAN KIAA0341.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB011124; BAA25478.1; -
KW Hypothetical protein; Coiled
FT DOMAIN 317 496 COILED COIL (POTENTIAL).
FT DOMAIN 571 639 COILED COIL (POTENTIAL).
FT DOMAIN 60 63 POLY-GLY.
FT DOMAIN 122 125 POLY-SER.
FT DOMAIN 282 290 POLY-SER.
FT DOMAIN 298 306 POLY-GLY.
FT DOMAIN 550 553 POLY-ALA.
SQ SEQUENCE 673 AA; 71790 MW; 3E548EC03A01F770 CRC64;

Query Match
Best Local Similarity 68.0%; Score 34; DB 1; Length 673;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RELATERESTS 11
DB 362 RELATELOGCS 372

RESULT 12
SYL_MYCTU STANDARD; PRT; 1041 AA.
AC 010765; 006181;
BT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Isolucyl-tRNA synthetase (EC 6.1.1.5) (Isolucine--tRNA ligase)
DE (ILERS).
CN ILRS OR RV1536 OR M11587 OR MYCY48.29C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_Taxid=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felkell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT *Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.*;
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolava M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

```



RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RL laboratory strains." to the EMBL/GenBank/DBJ databases.  
 CC Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) -> AMP +  
 CC diphosphate + L-isoleucyl-tRNA(Ile).  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
 CC -----  
 DR EMBL: 274020; CA98326.1; -  
 DR EMBL: AE007025; AA45854.1; -  
 DR HSRP: P56690; 11LE.  
 DR TIGR: MT1587; -  
 DR TubercuList; RV1536; -  
 DR InterPro: IPR002300; tRNA-synt\_1a.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002301; tRNA-synt\_1le.  
 DR Pfam: PF00133; tRNA-synt\_1; 1.  
 DR PRINTS: PR00984; TRNASYNTHTLE.  
 DR TIGRPFAMS: TIGR00392; 11es; 1.  
 DR PROSITE: PS00178; AA:TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Metal-binding; Zinc; Complete proteome.  
 FT SITE 53 63 "HIGH" REGION.  
 FT SITE 619 623 "KMSKS" REGION.  
 FT BINDING 622 622 ATP (BY SIMILARITY).  
 SQ SEQUENCE 1041 AA; 117339 MW; B5023822848E08C6 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 1041;  
 Best Local Similarity 70.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KETAELEST 10  
 DB 972 RELQELRKST 981

RESULT 13  
 YOBO\_BACSU  
 ID YOBO\_BACSU STANDARD: PRT; 1585 AA.  
 AC P45931;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yqbo.  
 GN YOBO.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 OX [1]  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / JH642;  
 RX MEDLINE=95219086; PubMed=7704261;  
 RA Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;  
 RT "Complete nucleotide sequence of a skin element excised by DNA  
 RT rearrangement during sporulation in Bacillus subtilis.";  
 RL Microbiology 141:323-327(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / JH642;  
 RX MEDLINE=97124195; PubMed=8969508;  
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,  
 RA Kobayashi Y.;  
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of

RT the Bacillus subtilis genome containing the skin element and many  
 RT sporulation genes.";  
 RL Microbiology 142:3103-3111(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara M., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Biotin A., Borchert S.,  
 RA Brolliet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,  
 RA Denzot F., Devine K.M., Dusterhoff A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,  
 RA Fritz C., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Gollighly E.J., Grandi G.,  
 RA Guisepi G., Guy B.J., Haga K., Hahsch J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hult M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Kjaer-blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takuchi M., Tanakoshi A., Tanaka T., Terpsira P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandendol M., Vannier F., Vassartotti A.,  
 RA Viari A., Wamuth R., Wedler E., Wedler H., Weltzenegger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasuno K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis".  
 RL Nature 390:249-256(1997).  
 RN [4]  
 RP IDENTIFICATION.  
 RX MEDLINE=96084975; PubMed=7489895;  
 RA Medigue C., Moszer I., Viari A., Danchin A.;  
 RT "Analysis of a Bacillus subtilis genome fragment using a co-operative  
 RT computer system prototype.";  
 RL Gene 165:GC37-GC51(1995).  
 CC -1- SIMILARITY: STRONG, TO B. SUBTILIS XKDO.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
 CC -----  
 DR EMBL: D32216; BAA06947.1; -  
 DR EMBL: D84432; BAA12411.1; -  
 DR EMBL: Z99117; CAB14544.1; -  
 DR Subtilist; BG11286; Yqbo.  
 DR InterPro: IPR00189; SLT\_domain.  
 DR Pfam: PF01464; SLT; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 1585 AA; 171030 MW; 1F82AB7F0250735F CRC64;

Query Match 68.0%; Score 34; DB 1; Length 1585;  
 Best Local Similarity 72.7%; Pred. No. 1.9e+02;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KETAELEST 11  
 DB 253 KALAELEST 263

```

RESULT 14
HEX9_ADE07 STANDARD: PRT: 138 AA.
ID HEX9_ADE07 STANDARD: PRT: 138 AA.
AC P03283:
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hexon-associated protein (Protein IX).
GN PIX.
OS Human adenovirus type 7, and
OC Vitruvius, dsDNA viruses, no RNA stage: Adenoviridae: Mastadenovirus.
CX NCBI_TaxID=10519, 45659;
[1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human adenovirus type 7; STRAIN=Gomen;
RX MEDLINE=81261948; PubMed=6266923;
RA Dijkema R., Maat J., Dekker B.M.M., van Ormondt H., Boyer H.W.;
RT "The gene for polypeptide IX of human adenovirus type 7."
RL Gene 13:375-385(1981).
[2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human adenovirus type 3;
RX MEDLINE=81261949; PubMed=7262560;
RA Engler J.A.;
RT "The nucleotide sequence of the polypeptide IX gene of human
  adenovirus type 3."
RL Gene 13:387-394(1981).
CC -1- FUNCTION: THIS PROTEIN IS A STRUCTURAL COMPONENT OF THE VIRION. IT
  MAY HAVE AN ADDITIONAL ROLE DURING ADENOVIRUS MULTIPLICATION.
CC -
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X03000; CAA26764.1; -
DR EMBL: J01962; AAA42510.1; -
DR PIR: A03854; SXAD97.
DR PIR: B03854; SXAD93.
KW Hexon-associated protein.
SQ SEQUENCE 138 AA; 14107 MW; B45EFBD6F7933EF2 CRC64;

Query March 66.0%; Score 33; DB 1; Length 138;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELAELREST 10
   1:1:11111
Db 119 KOVALREQT 128

RESULT 15
MT16_EMENT STANDARD: PRT: 293 AA.
ID MT16_EMENT STANDARD: PRT: 293 AA.
AC P56859:
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Phosphadenosine phosphosulfate reductase (EC 1.8.99.4) (PAPS
  reductase, thioresoxin dependent) (PADOPS reductase) (3'-
  phosphoadenylylsulfate reductase).
GN SA.
OS Emeritella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emeritella.
CX NCBI_TaxID=5072;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95287864; PubMed=7770049;

```

```

RA Borges-Walmsley M.I., Turner G., Bailey A.M., Brown J., Lehmbeck J.,
  Clausen I.G.;
RT "Isolation and characterisation of genes for sulphate activation and
  reduction in Aspergillus nidulans: implications for evolution of an
  RT allosteric control region by gene duplication."
RL Mol. Gen. Genet. 247:423-429(1995).
CC -1- FUNCTION: THE NADP DEPENDENT REDUCTION OF PAPS INTO SULFITE
  INVOLVES THIOREDOXIN WHICH PROBABLY PLAYS THE ROLE OF A THIOL
  CARRIER.
CC -1- CATALYTIC ACTIVITY: 5-phosphadenosine 3-phosphosulfate + reduced
  thioresoxin -> phosphadenosine phosphate + oxidized thioresoxin +
  sulfite.
CC -1- PATHWAY: THIRD STEP IN THE SULFATE ASSIMILATION PATHWAY THAT
  LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.
CC -1- SIMILARITY: BELONGS TO THE PAPS REDUCTASE FAMILY. CYSH SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X82555; -; NOT_ANNOTATED_CDS.
DR InterPro: IPR002500; PAPS_reduct.
DR InterPro: IPR004511; CYSH.
DR Pfam: PF01507; PAPS_reduct; 1.
DR TIGRFAMs: TIGR00434; CYSH; 1.
KW Methionine biosynthesis; Cysteine biosynthesis; Oxidoreductase; NADP.
SQ SEQUENCE 293 AA; 33978 MW; FCB8J9183F21D5F CRC64;

Query March 66.0%; Score 33; DB 1; Length 293;
Best Local Similarity 77.8%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ELAELRDST 10
   1:111111
Db 10 ETAEIRDST 18

```

Search completed: March 13, 2003, 18:49:50  
 Job time : 7.8 secs

GenCore version 5.1.4-P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:17 ; Search time 10.8 Seconds

(Without alignments)  
97.915 Million cell updates/sec

Title: US-10-050-200-3

Perfect score: 50

Sequence: 1 KEIAELRESTS 11

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.73:\*  
2: PIR1.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	281	2 138707	Fas ligand - human
2	41	82.0	278	2 A49266	Fas ligand - rat
3	41	82.0	279	2 A53062	Fas ligand - mouse
4	38	76.0	383	2 C69442	succinyl-CoA synth
5	37	74.0	408	1 MNMIVA	immediate-early pr
6	37	74.0	408	2 A72869	early gene transac
7	35	70.0	317	2 A59292	probable type II m
8	35	70.0	370	2 H64480	hypothetical prote
9	35	70.0	436	2 B83005	probable carboxyl-
10	35	70.0	625	2 E71297	probable flagellar
11	35	70.0	1333	2 S63403	probable membrane
12	34	68.0	74	2 H84256	hypothetical prote
13	34	68.0	159	2 C71080	hypothetical prote
14	34	68.0	294	2 T33588	hypothetical prote
15	34	68.0	304	2 T50721	systemin - chicken
16	34	68.0	328	2 T01944	hypothetical prote
17	34	68.0	568	2 E90364	hypothetical prote
18	34	68.0	673	2 T00328	hypothetical prote
19	34	68.0	1041	2 E70760	probable 11es prot
20	34	68.0	1585	2 B65948	phage-related prot
21	33	66.0	2228	1 T14029	variant-specific s
22	33	66.0	138	1 SKAD97	hexon-associated p
23	33	66.0	138	1 SKAD93	transcription regu
24	33	66.0	267	2 H75429	pyruvate dehydroge
25	33	66.0	270	2 A13642	sapc protein homol
26	33	66.0	356	2 F82247	solub protein, pept
27	33	66.0	374	2 T02652	probable choline k
28	33	66.0	387	2 E81399	succinate-CoA liga
29	33	66.0	387	2 E81399	

30	33	66.0	405	1 MNMVP1
31	33	66.0	405	2 T10420
32	33	66.0	413	2 T34117
33	33	66.0	464	2 T34332
34	33	66.0	473	2 F70031
35	33	66.0	491	2 F95010
36	33	66.0	494	2 E81846
37	33	66.0	494	2 G81095
38	33	66.0	499	2 A84171
39	33	66.0	514	2 C97882
40	33	66.0	576	2 B86499
41	33	66.0	576	2 B72125
42	33	66.0	583	2 C81546
43	33	66.0	1101	2 T31550
44	33	66.0	1162	2 A12843
45	33	66.0	1186	2 B97621

## ALIGNMENTS

RESULT 1  
138707  
Fas ligand - human  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence-revision 29-May-1998 #text-change 21-Jul-2000  
C:Accession: 138707; JC2340; S57565; 138554  
R:Takahashi, T.; Tanaka, M.; Inazawa, J.; Abe, T.; Suda, T.; Nagata, S.  
Int. Immunol. 6, 1567-1574, 1994  
A:Title: Human Fas ligand: gene structure, chromosomal location and species specific  
A:Reference number: 138707; MUID:95127560; PMID:7826947  
A:Accession: 138707  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-281 <RES>  
A:Cross-references: EMBL:U11821; NID:9595430; PIDN:AAC50124.1; PID:9595431  
R:Mita, E.; Hayashi, N.; Iio, S.; Takahara, T.; Hijioka, T.; Kasahara, A.; Fusamoto,  
Biochem. Biophys. Res. Commun. 204, 468-474, 1994  
A:Title: Role of Fas ligand in apoptosis induced by hepatitis C virus infection.  
A:Reference number: JC2340; MUID:95071350; PMID:7980502  
A:Accession: JC2340  
A:Molecule type: DNA  
A:Residues: 1-281 <MIT>  
A:Cross-references: GB:D38122; DDBJ:D29820; NID:9601892; PIDN:BA07320.1; PID:9136990  
R:Schatzlein, C.E.  
Submitted to the EMBL Data Library, June 1995  
A:Reference number: S57565  
A:Accession: S57565  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-281 <SCH>  
A:Cross-references: EMBL:X89102; NID:9887455; PID:9887456  
R:Allderson, M.R.; Tough, T.W.; Davis-Smith, T.; Braddy, S.; Falk, B.; Schooley, K.A.;  
J. Exp. Med. 181, 71-77, 1995  
A:Title: Fas ligand mediates activation-induced cell death in human T lymphocytes.  
A:Reference number: 138554; MUID:95105731; PMID:7528780  
A:Accession: 138554  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-281 <RE2>  
A:Cross-references: EMBL:U08137; NID:9624627; PIDN:AAC50071.1; PID:9624628  
C:Genetics:  
A:Gene: FasL  
A:Introns: 151/1; 116/3  
C:Keywords: glycoprotein; transmembrane protein  
F:80-102/Domain: transmembrane status predicted <TM>  
F:76,184,250,260/Binding site: carbohydrate (asn) (covalent) status predicted  
Query Match 100.0%; Score 50; DB 2; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.057;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|||||  
Db 109 KELAELRESTS 119

## RESULT 2

A49266

fas ligand - rat

C:Species: Rattus norvegicus (Norway rat)

C&gt;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999

C:Accession: A49266

R:Suda, T.; Takahashi, T.; Golstein, P.; Nagata, S.

Cell 75, 1169-1178, 1993

A:Title: Molecular cloning and expression of the Fas ligand, a novel member of the tumor

A:Reference number: A49266; MUID:94084792; PMID:7503205

A:Accession: A49266

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-278 &lt;SUD&gt;

A:Cross-references: GB:003470; NID:9440178; PIDN:AAC52129.1; PID:9440179

C:Keywords: glycoprotein; transmembrane protein

Query Match

Best Local Similarity 81.8%; Pred. No. 3;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11

Db 106 KELAELREFTN 116

## RESULT 3

A53062

fas ligand - mouse

C:Species: Mus musculus (house mouse)

C&gt;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999

C:Accession: A53062

R:Takahashi, T.; Tanaka, M.; Brannan, C.I.; Jenkins, N.A.; Copeland, N.G.; Suda, T.; Nag

Cell 76, 969-976, 1994

A:Title: Generalized lymphoproliferative disease in mice, caused by a point mutation in

A:Reference number: A53062; MUID:94185175; PMID:7511063

A:Accession: A53062

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-279 &lt;TAK&gt;

A:Cross-references: GB:006948; NID:9473564; PIDN:AAA17800.1; PID:9473565

Query Match

Best Local Similarity 82.0%; Score 41; DB 2; Length 279;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11

Db 107 KELAELREFTN 117

## RESULT 4

C69442

sucC1-1-CoA synthetase, beta subunit (sucC-1) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C&gt;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Jul-1999

C:Accession: C69442

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.F.

Glodex, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: C69442

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-383 &lt;KLE&gt;

A:Cross-references: GB:AE000996; GB:AE000782; NID:92689319; PIDN:AA89706.1; PID:9264

C:Superfamily: succinate-CoA ligase (ADP-forming) beta chain

Query Match

Best Local Similarity 76.0%; Score 38; DB 2; Length 383;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ELAELREST 10

Db 220 ELAELREFAT 228

## RESULT 5

MNAVIA

immediate-early protein IE-N - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

C&gt;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999

C:Accession: A39150

R:Carson, D.D.; Summers, M.D.; Guarino, L.A.

Virology 182, 279-286, 1991

A:Title: Molecular analysis of a baculovirus regulatory gene.

A:Reference number: A39150; MUID:91220660; PMID:2024466

A:Accession: A39150

A:Molecule type: DNA

A:Residues: 1-408 &lt;CAR&gt;

A:Cross-references: GB:M59422; NID:9332437; PIDN:AAA46701.1; PID:9332438

C:Superfamily: AcMNPV immediate-early protein IE-N

C:Keywords: DNA binding; early protein; tandem repeat; transcription regulation

F:34-49/Region: 7-residue repeats

F:51-58/Region: 4-residue repeats

F:190-196/Region: glutamine-rich

Query Match

Best Local Similarity 72.0%; Score 37; DB 1; Length 408;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11

Db 311 KELSRLAKTS 321

## RESULT 6

A72869

early gene transactivator - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

A:Note: dsDNA virus

C:Accession: A72869

R:Ayres, M.D.; Howard, S.C.; Kuzlo, J.; Lopez-Ferber, M.; Possee, R.D.

Virology 202, 586-605, 1994

A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis vir

A:Reference number: A72869; MUID:94303173; PMID:8030224

A:Accession: A72869

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-408 &lt;AYR&gt;

A:Cross-references: GB:I22858; NID:9510708; PIDN:AAA66781.1; PID:9559220

C:Genetics:

A:Gene: Ac-IE-2

C:Superfamily: AcMNPV immediate-early protein IE-N

Query Match

Best Local Similarity 74.0%; Score 37; DB 2; Length 408;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11

Db 311 KELSRLAKTS 321

## RESULT 7

A59292

probable type II myosin heavy chain - slime mold (Physarum polycephalum) (fragment)

```

C:Species: Physarum polycephalum
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 03-Nov-2000
C:Accession: A59292
R:Bailey, J.; Cook, L.J.; Kilme-Barber, R.; Swanton, E.; Solnic-Krezel, L.; Lohman, K.N.
submitted to GenBank, March 1999
A:Description: Identification of three genes expressed primarily during development in F
A:Reference number: A59292
A:Accession: A59292
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-317 <BAI>
A:Cross-references: GB:A1133501; NID:94494062; PIDN:CA839170.1; PID:94494063
A:Experimental source: strain Cl; dev stage plasmodium formation; clone 1b M8S; clone
A:Genetics:
A:Gene: mynd
C:Superfamily: kinetoplast-associated protein

Query Match 70.0%; Score 35; DB 2; Length 317;
Best Local Similarity 70.0%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KEIAELREST 10
:|||||:
Db 144 RELAELEKRS 153

RESULT 8
H64480
hypothetical protein MJ1449 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: H64480
R:Butl, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodok, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kalne, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: H64480
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-370 <BUU>
A:Cross-references: GB:U67585; GB:L77117; NID:91592088; PIDN:AAB99461.1; PID:91592094; T
A:Genetics:
A:Map position: FOR1418579-1419691

Query Match 70.0%; Score 35; DB 2; Length 370;
Best Local Similarity 72.7%; Pred. No. 58;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KEIAELRESTS 11
:|||||:
Db 126 KENKELREATS 136

RESULT 9
B83005
probable carboxyl-terminal proteinase PA5134 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83005
R:Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laribig, K.; Lim,
Lory, S.; Olson, M.V
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83005
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-436 <STO>
A:Cross-references: GB:AE004926; GB:AE004091; NID:99951424; PIDN:AG08519.1; GSPDB:GN001

```

```

A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5134
C:Superfamily: carboxyl-terminal processing proteinase

Query Match          70.0%: Score 35; DB 2; Length 436;
Best Local Similarity 63.6%: Pred. No. 69;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 KEIAELRESTS 11
        ::|||:||||
Db       91 EDPAELQESTS 101

RESULT 10
E71297
Probable flagellar hook-associated protein 1 (flgK) - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: E71297
R:Fraser, C.M.; Norris, S.J.; Welnsstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gerson, J.; Khalik, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utteback, T.; M. they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:965876
A:Accession: E71297
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-625 <COL>
A:Cross-references: GB:AE001240; GB:AE000520; MID:93322952; PIDN:AA65634.1; PID:9332
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0660

Query Match          70.0%: Score 35; DB 2; Length 625;
Best Local Similarity 54.5%: Pred. No. 1e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1 KEIAELRESTS 11
        |||::|::||
Db       574 KETIDMRDATS 584

RESULT 11
S63403
Probable membrane protein YNR070w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N3568
C:Species: Saccharomyces cerevisiae
C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 19-Apr-2002
C:Accession: S63403; S63402
R:Andree, B.; Iraqi, Housseini, I.; Urrestarazu, L.A.; Vissers, S.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S62920
A:Accession: S63403
A:Molecule type: DNA
A:Residues: 1-1333 <AND>
A:Cross-references: EMBL:Z71685; NID:91302603; PIDN:CAA96352.1; PID:e239601; PID:g130
A:Experimental source: strain S288C
R:Dusterhoeft, A.; Floeth, M.; Filtz, C.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D
submitted to the Protein Sequence Database, April 1996
A:Reference number: S62944
A:Accession: S63402
A:Molecule type: DNA
A:Residues: 1-1053 <DUE>
A:Cross-references: EMBL:Z71685; MIPS:YNR070w
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 14R
A:Cross-references: SGD:S0005353
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F:46-257/Domain: ATP-binding cassette homology <ABC1>

```

F:427-443/Domain: transmembrane #status predicted <TM1>  
 E:478-494/Domain: transmembrane #status predicted <TM2>  
 F:535-551/Domain: transmembrane #status predicted <TM3>  
 F:642-658/Domain: transmembrane #status predicted <TM4>  
 F:748-946/Domain: ATP-binding cassette homology <ABC2>  
 F:765-772/Region: nucleotide-binding motif A (P-loop)

Query Match 70.0%; Score 35; DB 2; Length 1333;  
 Best Local Similarity 72.7%; Pred. No. 2.3e+02;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 KELAELRESTS 11  
 DB 1122 KELEYEVRKESK 1132

RESULT 12  
 H84256  
 Hypothetical protein Vng1003h [Imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: H84256  
 R:Ng, W.V.; Kennedy, S.P.; Maharras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Lelthausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jaido Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A:Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: H84256  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-74 <STO>  
 A:Cross-references: GB:AE004437; NID:g10580560; PIDN:AMG19420.1; GSPDB:GN00138  
 A:Gene: VNG1003H

Query Match 68.0%; Score 34; DB 2; Length 74;  
 Best Local Similarity 72.7%; Pred. No. 16;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 KELAELRESTS 11  
 DB 15 KELAELVESTS 25

RESULT 13  
 C71080  
 Hypothetical protein PH0907 - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: C71080  
 R:Kavaraevskiy, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hiro, Y.; Yamamoto, S.; Sekit M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kuhlida, N.; Oguchi DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic A A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: C71080  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-159 <KAW>  
 A:Cross-references: GB:AP000004; NID:g3236131; PIDN:BA30001.1; PID:g3257318  
 A:Experimental source: strain OT3  
 A>Note: this accession replaces an interim accession for a sequence replaced by GenBank C:Genetics:  
 A:Gene: PH0907  
 C:Superfamily: hypothetical protein MDEC503

Query Match 68.0%; Score 34; DB 2; Length 159;  
 Best Local Similarity 63.6%; Pred. No. 36;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 KELAELRESTS 11

DB 149 KEYEERKSTKS 159

RESULT 14  
 T33586

Hypothetical protein C40A11.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T33586  
 R:Maggi, L.; Goela, D.  
 submitted to the EMBL Data Library, October 1998  
 A:Description: The sequence of C. elegans cosmid C40A11.  
 A:Reference number: 221374  
 A:Accession: T33586  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-294 <MAC>  
 A:Cross-references: EMBL:AF099914; PIDN:AMC68758.1; GSPDB:GN00020; CESP:C40A11.4  
 A:Experimental source: strain Bristol N2; clone C40A11  
 C:Genetics:  
 A:Gene: CESP:C40A11.4  
 A:Map position: 2  
 A:Introns: 33/2; 74/3; 103/3; 183/3; 222/1

Query Match 68.0%; Score 34; DB 2; Length 294;  
 Best Local Similarity 87.5%; Pred. No. 70;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KELAELRE 8  
 DB 138 KELEYEVR 145

RESULT 15.  
 I50721  
 synein - chicken (fragment)  
 C:Species: Gallus gallus (chicken)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 03-Nov-2000  
 C:Accession: I50721  
 R:Becker, B.; Bellin, R.M.; Sernet, S.W.; Huilatt, T.W.; Robson, R.M.  
 Biochem. Biophys. Res. Commun. 213, 796-802, 1995  
 A:Title: Synein contains the rod domain of intermediate filaments.  
 A:Reference number: I50721; MUID:95382823; PMID:7654240  
 A:Accession: I50721  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-304 <BEC>  
 A:Cross-references: EMBL:U28143; NID:g1052880; PID:g1052881  
 C:Superfamily: cytoskeletal keratin

Query Match 68.0%; Score 34; DB 2; Length 304;  
 Best Local Similarity 87.5%; Pred. No. 73;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KELAELRE 8  
 DB 31 KELAELRE 38

Search completed: March 13, 2003, 18:52:55  
 Job time: 13.8 secs

Non-bleavable Fas  
Fas ligand (Fast)  
Fas ligand (Fast)  
Human Fas-L protein  
Human Fas ligand  
Human Fas ligand d  
Human Fas ligand  
Human Fas ligand  
Fas ligand  
Wild type Fas liga  
Human Fas ligand d  
Human Fas ligand  
Human FasL protein  
Amino acid sequenc  
Human Fas ligand  
Human Fas ligand  
Human Fas ligand  
Human Fas ligand  
Human Fas ligand  
Human Fas ligand  
Human Fas ligand p  
Human Fas ligand p  
Human Fas ligand p  
CH3/Fas<sup>+</sup>ligand dom  
IG2G CH3/Fas<sup>+</sup>dom  
Secreted modified  
Human Fas ligand  
Mouse Fas ligand  
Rat Fas ligand (pa  
Rat Fas ligand enc

DR WPI: 1995-194031/25.  
 XX Peptide which binds to Fas antigen, and antibody reactive with it -  
 PT for treatment and diagnosis of viral or auto-immune diseases  
 XX  
 PS Example 21: Page 196-197; 300pp; Japanese.  
 XX  
 CC CD179 is a truncated human Fas ligand from which the Leu residue at  
 CC the C-terminus (position 179) has been deleted. Fas ligands able to  
 CC induce apoptosis in cells which express the Fas cell surface antigen  
 CC are claimed.  
 XX  
 SQ Sequence 178 AA:  
 Query Match 100.0%; Score 50; DB 16; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 0.054;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KELAELRESTS 11  
 |||||  
 Db 7 KELAELRESTS 17  
 RESULT 2  
 AAR79099  
 ID AAR79099 standard; Protein; 179 AA.  
 XX  
 AC AAR79099;  
 XX  
 DT 21-FEB-1996 (first entry)  
 XX  
 DE Human Fas ligand (partial sequence).  
 XX  
 KW Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;  
 KW Fas cell surface antigen; Fas-L; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9513293-A1.  
 XX  
 PD 18-MAY-1995.  
 XX  
 PF 10-NOV-1994; 94WO-JP01899.  
 XX  
 PR 18-OCT-1994; 94JP-0278378.  
 PR 10-NOV-1993; 93JP-0305975.  
 PR 13-DEC-1993; 93JP-0342526.  
 PR 18-MAR-1994; 94JP-0074344.  
 PR 08-JUL-1994; 94JP-0180953.  
 PR 07-SEP-1994; 94JP-0239363.  
 XX  
 PA (MOCH) MOCHIDA PHARM CO LTD.  
 PA (OSAB-) OSAKA BIOSCIENCE INST.  
 XX  
 PI Nagata S, Nakamura N, Suda T, Takahashi T;  
 XX  
 DR WPI: 1995-194031/25.  
 DR N-PSDB; AAO94157.  
 XX  
 PT Peptide which binds to Fas antigen, and antibody reactive with it -  
 PT for treatment and diagnosis of viral or auto-immune diseases  
 XX  
 PS Claim 3: Page 222-224; 300pp; Japanese.  
 XX  
 CC Fas ligands or active fragments able to induce apoptosis in cells  
 CC which express the Fas cell surface antigen are claimed. The  
 CC proteins are isolated from human, rat and mouse sources. The present  
 CC sequence represents part of the human Fas ligand.  
 XX  
 SQ Sequence 179 AA:  
 Query Match 100.0%; Score 50; DB 16; Length 179;  
 Best Local Similarity 100.0%; Pred. No. 0.054;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KELAELRESTS 11  
 |||||  
 Db 7 KELAELRESTS 17  
 RESULT 3  
 AAM11814  
 ID AAM11814 standard; peptide; 179 AA.  
 XX  
 AC AAM11814;  
 XX  
 DT 20-OCT-1997 (first entry)  
 XX  
 DE Fas ligand.  
 XX  
 KW Human; Fas ligand; antigen; neutral; antibody; apoptosis; HIV;  
 KW induction; assay; enzyme linked immunosorbent assay; diagnosis;  
 KW disease; hepatitis B; hepatitis C; human immunodeficiency virus;  
 KW graft versus host disease; ulcerative colitis; sequelae;  
 KW myocardial infarction; treatment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9702290-A1.  
 XX  
 PD 23-JAN-1997.  
 XX  
 PF 01-JUL-1996; 96WO-JP01820.  
 PR 17-MAY-1996; 96US-0649100.  
 PR 30-JUN-1995; 95JP-0188480.  
 XX  
 PA (MOCH) MOCHIDA PHARM CO LTD.  
 PA (OSAB-) OSAKA BIOSCIENCE INST.  
 XX  
 PI Co MS, Matsusue T, Nagata S, Shirakawa K, Vasquez M;  
 XX  
 DR WPI: 1997-108917/10.  
 XX  
 PT Antibody reactive with Fas ligand capable of inducing apoptosis -  
 PT used for diagnostic assay of Fas ligand in body fluids and for  
 PT treatment of diseases in which Fas ligand/Fas antigen is involved  
 XX  
 PS Disclosure: Fig 1; 164pp; Japanese.  
 XX  
 CC Antigens derived from the present human Fas ligand sequence can  
 CC be used in the preparation of a neutral antibody against an  
 CC apoptosis inducing Fas ligand. The antibody can be used in a Fas  
 CC ligand assay, e.g. an enzyme linked immunosorbent assay, to  
 CC diagnose diseases in which the Fas ligand/Fas antigen system is  
 CC implicated, e.g. hepatitis B/C, human immunodeficiency virus,  
 CC graft/host disorders, ulcerative colitis or sequelae of myocardial  
 CC infarction. The antibody may also be used to treat such diseases.



```
XX Sequence 179 AA;
Query Match 100.0%; Score 50; DB 18; Length 179;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KEIAELRESTS 11
   |||
Db 7 KEIAELRESTS 17

RESULT 4
AAW10875
AAW10875 standard; protein; 180 AA.
XX
XX AAW10875;
XX
XX 25-SEP-1997 (first entry)
XX
XX Fasl/LT-alpha hybrid.
XX DE
XX C-terminal extracellular domain; CD40 ligand; mutated; trimerisation;
XX lymphotoxin; LT; beta; alpha; tumour necrosis factor; TNF; inhibitor;
XX signalling; receptor; Fas; treating; autoimmune disease; diabetes;
XX rheumatoid arthritis; inflammatory bowel disease; HIV proliferation;
XX immune response; tissue graft.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 44 /note- "LT-alpha subunit association domain residue"
XX FT Misc-difference 68..70 /note- "LT-alpha subunit association domain residues"
XX FT Misc-difference 70 /note- "given as O in the specification"
XX FT Misc-difference 92 /note- "LT-alpha subunit association domain residue"
XX FT Misc-difference 94 /note- "LT-alpha subunit association domain residue"
XX FT Misc-difference 122..127 /note- "LT-alpha subunit association domain residue"
XX FT Misc-difference 126 /note- "LT-alpha subunit association domain residues"
XX FT Misc-difference 126 /note- "given as O in the specification"
XX FT Misc-difference 129..130 /note- "LT-alpha subunit association domain residues"
XX FT Misc-difference 139 /note- "LT-alpha subunit association domain residue"
XX FT Misc-difference 146 /note- "LT-alpha subunit association domain residue"
XX FT Misc-difference 172 /note- "LT-alpha subunit association domain residue"
XX FT Misc-difference 177..178 /note- "LT-alpha subunit association domain residues"
XX FT Misc-difference 177..178 /note- "LT-alpha subunit association domain residues"
XX
XX WO9640774-A1.
XX
XX 19-DEC-1996.
XX
XX 06-JUN-1996; 96NO-US09773.
XX
XX 07-JUN-1995; 95US-0476074.
XX
XX (BIOC ) BIOGEN INC.
XX
XX Browning JL, Karpusas MN, Meier W;
XX WPI; 1997-087055/08.
XX
XX Heteromeric complexes comprising lymphotoxin sub-units - useful as
XX inhibitors of signalling by TNF- and TNF-related receptors, for
XX treating autoimmune disease, graft rejection, etc.
```

```
XX PS Disclosure; Fig 10; 87pp; English.
XX
XX This sequence comprises the C-terminal extracellular domain of the Fas
XX ligand altered to drive trimerisation with wild-type lymphotoxin (LT)-
XX beta. LT subunit association domains potentiate subunit associations into
XX an active trimeric ligand. Tumour necrosis factor (TNF)-related ligand
XX monomers mutated in their respective subunit associations domains can
XX form heteromeric complexes with LT subunits. LT-alpha/beta heteromeric
XX complexes and LT heterotrimeric complexes are useful as inhibitors of
XX signalling by TNF receptors (R) and TNF-related receptors, such as CD40,
XX Fas and LT-beta receptors. By inhibiting TNF-R signalling the complexes
XX (esp., LT-alpha2/beta1) can be used for treating autoimmune diseases
XX including diabetes, rheumatoid arthritis, inflammatory bowel disease and
XX HIV proliferation, known to be improved by TNF inhibition. The complexes
XX are also useful for preventing or reducing the severity of an immune
XX response to a tissue graft.
XX
XX Sequence 180 AA;
Query Match 100.0%; Score 50; DB 18; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KEIAELRESTS 11
   |||
Db 7 KEIAELRESTS 17

RESULT 5
AAW10876
AAW10876 standard; protein; 180 AA.
XX
XX AAW10876;
XX
XX 25-SEP-1997 (first entry)
XX
XX Fasl/LT-beta hybrid.
XX DE
XX C-terminal extracellular domain; CD40 ligand; mutated; trimerisation;
XX lymphotoxin; LT; beta; alpha; tumour necrosis factor; TNF; inhibitor;
XX signalling; receptor; Fas; treating; autoimmune disease; diabetes;
XX rheumatoid arthritis; inflammatory bowel disease; HIV proliferation;
XX immune response; tissue graft.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 44 /note- "LT-beta subunit association domain residue"
XX FT Misc-difference 68..70 /note- "LT-beta subunit association domain residues"
XX FT Misc-difference 88..89 /note- "LT-beta subunit association domain residues"
XX FT Misc-difference 92 /note- "LT-beta subunit association domain residues"
XX FT Misc-difference 94 /note- "LT-beta subunit association domain residue"
XX FT Misc-difference 124..127 /note- "LT-beta subunit association domain residue"
XX FT Misc-difference 129..130 /note- "LT-beta subunit association domain residues"
XX FT Misc-difference 139 /note- "LT-beta subunit association domain residue"
XX FT Misc-difference 142 /note- "LT-beta subunit association domain residue"
XX FT Misc-difference 145..146 /note- "LT-beta subunit association domain residue"
XX FT Misc-difference 171..172 /note- "LT-alpha subunit association domain residues"
XX FT Misc-difference 177..178 /note- "LT-beta subunit association domain residues"
XX FT Misc-difference 177..178 /note- "LT-beta subunit association domain residues"
```

PN WO9640774-A1.  
 XX  
 XX 19-DEC-1996.  
 PD  
 XX  
 XX 06-JUN-1996; 96WO-US09773.  
 PF  
 XX 07-JUN-1995; 95US-0476074.  
 PR  
 XX (BIOJ ) BIOGEN INC.  
 PA  
 XX Browning JL, Karpusas MN, Meier W;  
 PI WPI: 1997-087055/08.  
 DR  
 XX Heteromeric complexes comprising lymphotoxin sub-units - useful as  
 PT inhibitors of signalling by TNF- and TNF-related receptors, for  
 PT treating auto-immune disease, graft rejection, etc.  
 XX  
 XX Disclosure: Fig 10; 87pp; English.  
 PS  
 XX This sequence comprises the C-terminal extracellular domain of the Fas  
 CC ligand altered to drive trimerisation with wild-type lymphotoxin (LT).  
 CC alpha. LT subunit association domains potentiate subunit associations to  
 CC an active trimeric ligand. Tumour necrosis factor (TNF)-related ligand  
 CC monomers mutated in their respective subunit associations domains can  
 CC form heteromeric complexes with LT subunits. LT-alpha/Beta heteromeric  
 CC complexes and LT heterotrimeric complexes are useful as inhibitors of  
 CC signalling by TNF receptors (R) and TNF-related receptors, such as CD40,  
 CC FAS and LT-beta receptors. By inhibiting TNF-R signalling the complexes  
 CC (esp., LT-alpha2/beta1) can be used for treating autoimmune diseases  
 CC including diabetes, rheumatoid arthritis, inflammatory bowel disease and  
 CC HIV proliferation, known to be improved by TNF inhibition. The complexes  
 CC are also useful for preventing or reducing the severity of an immune  
 CC response to a tissue graft.  
 XX  
 XX Sequence 180 AA:  
 SQ  
 Query Match 100.0%; Score 50; DB 18; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 0.055;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KELAELRESTS 11  
 DB 7 KELAELRESTS 17  
 RESULT 6  
 AAB86306  
 ID AAB86306 standard; protein; 213 AA.  
 XX  
 XX AAB86306;  
 AC  
 XX  
 XX 13-SEP-2001 (first entry)  
 DT  
 XX  
 XX Fast fusion construct FasL-hexamer.  
 DE  
 XX  
 XX Fusion protein; dimerization; oligomerization; C1q; vaccine; diphtheria;  
 KW collection; antiinflammatory; immunomodulatory; antiapoptotic; apoptotic;  
 KW antiviral; antibacterial; antimalarial; antiprotococcal; antitumor; BCG;  
 KW endocrinological; medicine; autoimmune disease; hyper-apoptotic state;  
 KW hypo-apoptotic state; infection; tumor; endocrinological disorder;  
 KW immunization; rubella; measles; poliomyelitis; rabies; tetanus; malaria;  
 KW yellow fever; human immune deficiency virus; Influenza.  
 KM  
 XX  
 XX OS Unidentified.  
 XX  
 XX DE19963859-A1.  
 FM  
 XX 12-JUL-2001.  
 PD  
 XX 30-DEC-1999; 99DE-1063859.  
 PF  
 XX 30-DEC-1999; 99DE-1063859.  
 PR

XX  
 PA (APOT-) APOTECHE RES & DEV LTD.  
 XX  
 XX Tschopp J, Schneider P, Holler N;  
 PI  
 XX WPI: 2001-426647/46.  
 DR  
 XX  
 XX New dimers or oligomers of oligomeric recombinant fusion proteins,  
 PT useful as pharmaceuticals and vaccines, comprises functional component  
 PT and component that induces oligomerization -  
 XX  
 XX Example 1; Page 17-18; 42pp; German.  
 PS  
 XX This invention describes novel dimers or oligomers (I) of dl-, tri-,  
 CC tetra- or penta-mers (II) of recombinant fusion proteins (FP) comprising  
 CC at least one each of components A and B, where A is a biologically active  
 CC protein (or fragment) and B is a protein (or fragment) that causes  
 CC dimerization/oligomerization of (I) without requiring a third molecule.  
 CC Particularly A functions as a ligand for antibodies, soluble or membrane  
 CC bound signalling molecules or receptors or is an antibody (or its  
 CC fragment). The invention also describes (a) FP containing as component B  
 CC a protein of the C1q or collectin families; (b) DNA sequences (II) that  
 CC encode the FP of (a); (c) expression vectors containing (II); (d) host  
 CC cells containing the vector of (c); and (e) (II) in which component B of  
 CC FP is derived from the immunoglobulin, C1q or collectin families of  
 CC proteins. The products of the invention have antiinflammatory,  
 CC immunomodulatory, antiapoptotic, apoptotic, antiviral, antibacterial,  
 CC antimalarial, antiprotococcal, antitumor and endocrinological activity. (I)  
 CC are useful in human or veterinary medicine, to: (i) produce  
 CC pharmaceuticals for treatment of inflammation; autoimmune diseases;  
 CC hyper- or hypo-apoptotic states; infections (particularly viral); tumors  
 CC (particularly of the lymphatic system) and/or endocrinological disorders;  
 CC (ii) produce vaccines for active or passive immunization against  
 CC infections (specifically rubella, measles, poliomyelitis, rabies,  
 CC tetanus, diphtheria, BCG, malaria, yellow fever, human immune deficiency  
 CC virus or Influenza), also as vaccine adjuvants; or (iii) for in vitro  
 CC diagnosis. Crosslinking of component A is now achieved without requiring  
 CC a third molecule (e.g. antibody), resulting in increased biological  
 CC activity and regeneration of activity. This sequence represents the FasL  
 CC trimer construct described in the method of the invention.  
 XX  
 XX Sequence 213 AA:  
 SQ  
 Query Match 100.0%; Score 50; DB 22; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 0.066;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KELAELRESTS 11  
 DB 41 KELAELRESTS 51  
 RESULT 7  
 AAB86307  
 ID AAB86307 standard; protein; 213 AA.  
 XX  
 XX AAB86307;  
 AC  
 XX  
 XX 13-SEP-2001 (first entry)  
 DT  
 XX  
 XX Fast fusion construct super-FasL.  
 DE  
 XX  
 XX Fusion protein; dimerization; oligomerization; C1q; vaccine; diphtheria;  
 KW collection; antiinflammatory; immunomodulatory; antiapoptotic; apoptotic;  
 KW antiviral; antibacterial; antimalarial; antiprotococcal; antitumor; BCG;  
 KW endocrinological; medicine; autoimmune disease; hyper-apoptotic state;  
 KW hypo-apoptotic state; infection; tumor; endocrinological disorder;  
 KW immunization; rubella; measles; poliomyelitis; rabies; tetanus; malaria;  
 KW yellow fever; human immune deficiency virus; Influenza; super-FasL.  
 KM  
 XX  
 XX OS Unidentified.  
 XX  
 XX DE19963859-A1.  
 FM

XX 12-JUL-2001.  
 PD 30-DEC-1999; 99DE-1063859.  
 PF 30-DEC-1999; 99DE-1063859.  
 PR 30-DEC-1999; 99DE-1063859.  
 XX (APOT-) APOTEC RES & DEV LTD.  
 PA  
 PI Tschopp J, Schneider P, Holler N;  
 XX WPI; 2001-426647/46.  
 DR  
 XX New dimers or oligomers of oligomeric recombinant fusion proteins,  
 PT useful as pharmaceuticals and vaccines, comprises functional component  
 PT and component that induces oligomerization -  
 XX  
 XX Example 2: Page 18-19; 42pp; German.  
 PS  
 XX This invention describes novel dimers or oligomers (I) of di-, tri-,  
 CC tetra- or penta-mers (II) of recombinant fusion proteins (FP) comprising  
 CC at least one each of components A and B, where A is a biologically active  
 CC protein (or fragment) and B is a protein (or fragment) that causes  
 CC dimerization/oligomerization of (I) without requiring a third molecule.  
 CC Particularly A functions as a ligand for antibodies, soluble or membrane  
 CC bound signalling molecules or receptors or is an antibody (or its  
 CC fragment). The invention also describes (a) FP containing as component B  
 CC a protein of the C1q or collectin families; (b) DNA sequences (II) that  
 CC encode the FP of (a); (c) expression vectors containing (III); (d) host  
 CC cells containing the vector of (c); and (e) (II) in which component B of  
 CC FP is derived from the immunoglobulin, C1q or collectin families of  
 CC proteins. The products of the invention have antiinflammatory,  
 CC immunomodulatory, antiapoptotic, apoptotic, antiviral, antibacterial,  
 CC antimalarial, antiparasitic, antitumor and endocrinological activity. (I)  
 CC are useful, in human or veterinary medicine, to: (1) produce  
 CC pharmaceuticals for treatment of inflammation; autoimmune diseases;  
 CC hyper- or hypo-apoptotic states; infections (particularly viral); tumors  
 CC (particularly of the lymphatic system) and/or endocrinological disorders;  
 CC (II) produce vaccines for active or passive immunization against  
 CC infections (specifically rubella, measles, poliomyelitis, rabies,  
 CC tetanus, diphtheria, BCG, malaria, yellow fever, human immune deficiency  
 CC virus or influenza), also as vaccine adjuvants; or (III) for in vitro  
 CC diagnosis. Crosslinking of component A is now achieved without requiring  
 CC a third molecule (e.g. antibody), resulting in increased biological  
 CC activity and regeneration of activity. This sequence represents the  
 CC super-FasL construct described in the method of the invention.  
 XX  
 XX SQ Sequence 213 AA;  
 XX  
 XX Query Match 100.0%; Score 50; DB 22; Length 213;  
 XX Best Local Similarity 100.0%; Pred. No. 0.066; 0; Indels 0; Gaps 0;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KELAELRESTS 11  
 DB 41 KELAELRESTS 51  
 XX  
 XX RESULT 8  
 XX ABB05003  
 XX ID ABB05003 standard; Protein: 238 AA.  
 XX  
 XX AC ABB05003;  
 XX  
 XX 20-MAR-2002 (first entry)  
 DT  
 XX FLAG peptide-leucine zipper-human Fas ligand fusion protein.  
 DE  
 XX Human; Fas ligand; Fas ligand-fusion protein; leucine zipper;  
 KW transmembrane protein; hepatotropic; immunosuppressive; apoptosis;  
 KW liver cell necrosis; viral hepatitis; autoimmune disease.  
 XX  
 OS Homo sapiens.

OS Synthetic.  
 XX  
 XX MO200190382-A2.  
 PN  
 XX 29-NOV-2001.  
 PD  
 XX 28-MAY-2001; 2001WO-JP04456.  
 PF  
 XX 26-MAY-2000; 2000JP-0156209.  
 PR  
 XX (MOCH ) MOCHIDA PHARM CO LTD.  
 PA  
 XX Touna J;  
 XX  
 XX WPI: 2002-097655/13.  
 DR  
 XX N-PSDB: ABA92524.  
 DR  
 XX Fusion proteins of a Fas ligand with an oligomer forming peptide and an  
 PT expression promoter peptide, useful for inducing apoptosis in cells  
 PT expressing Fas -  
 XX  
 XX Claim 7: Page 46-48; 48pp; Japanese.  
 PS  
 XX The present invention describes a fusion protein comprising a peptide  
 CC which contains all or part of the Fas ligand sequence, a peptide capable  
 CC of forming an oligomer and a peptide which increases the production  
 CC yield of recombinant protein. The fusion protein contains: (1) a FLAG  
 CC peptide sequence; (2) an leucine zipper sequence; and (3) the  
 CC extracellular domain of human Fas ligand. The fusion protein has  
 CC hepatotropic and immunosuppressive activities. Binding of the Fas ligand  
 CC sequence to Fas on the surface of cells expressing Fas induces apoptosis  
 CC in the cells. The fusion protein can be used in the treatment of  
 CC diseases with which Fas-expressing cells are associated, such as liver  
 CC cell necrosis produced by viral hepatitis, and autoimmune diseases.  
 CC The present sequence represents a fusion protein of the present  
 CC invention.  
 XX  
 XX SQ Sequence 238 AA;  
 XX  
 XX Query Match 100.0%; Score 50; DB 23; Length 238;  
 XX Best Local Similarity 100.0%; Pred. No. 0.074;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KELAELRESTS 11  
 DB 66 KELAELRESTS 76  
 XX  
 XX RESULT 9  
 XX AAY28597  
 XX ID AAY28597 standard; Protein: 261 AA.  
 XX  
 XX AC AAY28597;  
 XX  
 XX 26-NOV-1999 (first entry)  
 DT  
 XX Fas ligand (FasL) mutant delta 3.  
 DE  
 XX Fas ligand; FasL; apoptosis; non-cleavable; graft intolerance;  
 KW autoimmune destruction; cleavage site; mutleln.  
 KW  
 XX Homo sapiens.  
 OS  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 125..126  
 FT /note="Deletion of 20 wild-type amino acids (126-145)."  
 FT  
 XX MO9936079-A1.  
 PN  
 XX 22-JUL-1999.  
 PD  
 XX 07-JAN-1999; 99WO-US00667.  
 PF

XX 14-JAN-1998; 98US-0006755.  
 XX (REGC ) UNIV CALIFORNIA.  
 PA Backeskov S, Braat D, Kang S, Stock PG;  
 XX WPI; 1999-468942/39.  
 XX A non-cleavable Fas ligand polypeptide that has capacity to activate  
 PT Fas receptor-mediated apoptosis  
 PS  
 XX Claim 14: Page -: 81pp; English.  
 CC This sequence is Fas ligand (FasL) mutant delta 3. This protein was made  
 CC by deleting amino acids 126-145 (inclusive) of the wild-type FasL  
 CC sequence AAY28594. The FasL protein activates the Fas receptor which  
 CC induces apoptosis. The FasL deletion mutants (AAY28595-Y28597) are  
 CC non-cleavable. The mutations inhibit the proteolytic cleavage of FasL  
 CC and have the capacity to activate a Fas receptor-mediated pathway.  
 CC Claimed mutants of FasL or an organ or tissue expressing a mutant can be  
 CC used to alleviate symptoms of a disorder characterized by inadequate or  
 CC inappropriate stimulation of a Fas receptor-mediated pathway, such as  
 CC apoptosis, in a tissue or organ. In particular, FasL mutants can be used  
 CC to treat intolerance to a graft. In a patient, FasL mutants protect  
 CC an organ or tissue from autoimmune destruction.  
 CC Note: The present sequence is not shown in the specification but it  
 CC has been derived from the wild-type FasL sequence AAY28594 shown in the  
 CC specification.  
 XX

SO Sequence 261 AA;

Query Match 100.0%; Score 50; DB 20; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 0.082;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
 |||||  
 DB 109 KELAELRESTS 119

RESULT 10  
 AAW48954  
 ID AAW48954 standard; Protein; 265 AA.  
 XX  
 AC AAW48954;

DT 23-SEP-1998 (first entry)

DE Non-cleavable Fas ligand 1306145 deletion muten.

XX Non-cleavable Fas ligand 1306145 deletion muten; Fas; erythematous;  
 KW gene therapy; autoimmune disease; multiple sclerosis;  
 KW rheumatoid arthritis; myasthenia gravis; transplant rejection;  
 KW glomerulonephritis.

OS Homo sapiens.  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH 1..129  
 FT /note= "represents residues 1-129 of the wild-type  
 FT Fas ligand"  
 FT Region 130..265  
 FT /note= "represents residues 146-281 of the  
 FT wild-type Fas ligand"

PN W09821232-A2.

PD 22-MAY-1998.

PP 13-NOV-1997; 97WO-US20864.

XX

PR 12-NOV-1997; 97US-0968686.  
 PR 13-NOV-1996; 96US-0030871.  
 PR 10-FEB-1997; 97US-0039972.

PA (CHIR ) CHIRON CORP.

PI Chu K;

DR WPI; 1998-297861/26.

DR N-PSDB; AAV32623.

PT New DNA encoding Fas ligand agonist including, e.g. deletion  
 PT useful for, e.g. treating auto-immune diseases or transplant  
 PT rejection

PS Claim 4: Pages 65-66; 72pp; English.

XX The present sequence represents a non-cleavable Fas ligand 1306145  
 CC deletion muten. Fas ligand deletion muten can be expressed in  
 CC cells transfected with the DNA (AAV32623) coding for the mutant protein.  
 CC These cells, expressing the mutant Fas ligand in a non-cleavable form,  
 CC are claimed to be useful in vitro to identify cells that express Fas  
 CC and, in vivo or in vitro, for reducing proliferation of Fas-expressing  
 CC cells. The DNA encoding the Fas ligand mutant is claimed to be  
 CC useful in gene therapy procedures and for the treatment of autoimmune  
 CC diseases, e.g. multiple sclerosis, erythematous, rheumatoid arthritis,  
 CC glomerulonephritis, myasthenia gravis and transplant rejection.

SO Sequence 265 AA;

Query Match 100.0%; Score 50; DB 19; Length 265;  
 Best Local Similarity 100.0%; Pred. No. 0.084;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
 |||||  
 DB 109 KELAELRESTS 119

RESULT 11  
 AAW48953  
 ID AAW48953 standard; Protein; 268 AA.  
 XX  
 AC AAW48953;

DT 23-SEP-1998 (first entry)

DE Non-cleavable Fas ligand 1306142 deletion muten.

XX Non-cleavable Fas ligand 1306142 deletion muten; Fas; erythematous;  
 KW gene therapy; autoimmune disease; multiple sclerosis;  
 KW rheumatoid arthritis; myasthenia gravis; transplant rejection;  
 KW glomerulonephritis.

OS Homo sapiens.  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH 1..129  
 FT /note= "represents residues 1-129 of the wild-type  
 FT Fas ligand"  
 FT Region 130..268  
 FT /note= "represents residues 143-281 of the  
 FT wild-type Fas ligand"

PN W09821232-A2.

PD 22-MAY-1998.

PP 13-NOV-1997; 97WO-US20864.

PR 12-NOV-1997; 97US-0968686.

PR 13-NOV-1996; 96US-0030871.

```

PR 10-FEB-1997; 97US-0039972.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Chu K;
XX
XX WPI; 1998-297861/26.
XX N-PSDB; AAV32622.
XX
XX New DNA encoding Fas ligand agonist including, e.g. deletion -
PT useful for, e.g. treating auto-immune diseases or transplant
PT rejection
XX
XX Claim 4; Pages 62-63; 72pp; English.
XX
XX The present sequence represents a non-cleavable Fas ligand 1306142
CC deletion mutein. Fas ligand deletion mutein can be expressed in
CC cells transfected with the DNA (AAV32622) coding for the mutant protein.
CC These cells, expressing the mutant Fas ligand in a non-cleavable form,
CC are claimed to be useful in vitro to identify cells that express Fas
CC and, in vivo or in vitro, for reducing proliferation of Fas-expressing
CC cells. The DNA encoding the Fas ligand mutant is claimed to be
CC useful in gene therapy procedures and for the treatment of autoimmune
CC diseases, e.g. multiple sclerosis, erythematosis, rheumatoid arthritis,
CC glomerulonephritis, myasthenia gravis and transplant rejection.
XX
XX Sequence 268 AA:
SQ
Query Match 100.0%; Score 50; DB 19; Length 268;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KEIAELRESTS 11
DB 109 KEIAELRESTS 119
RESULT 12
AAV28596
ID AAV28596 standard; Protein; 271 AA.
XX
XX AAV28596;
XX
XX 26-NOV-1999 (first entry)
XX
XX Fas ligand (FasL) mutant delta 2.
XX
XX Fas ligand; FasL; apoptosis; non-cleavable; graft intolerance;
XX autoimmune destruction; cleavage site; mutein.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 125..126 /note="Deletion of 10 wild-type amino acids (126-135)"
XX
XX WO9936079-A1.
XX
XX 22-JUL-1999.
XX
XX 07-JAN-1999; 99WO-US00667.
XX
XX 14-JAN-1998; 98US-0006755.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Baekkeskov S, Braat D, Kang S, Stock PG;
XX
XX WPI; 1999-468942/39.
XX
XX A non-cleavable Fas ligand polypeptide that has capacity to activate
XX Fas receptor-mediated apoptosis
PT

```

```

XX
XX Claim 13; Page -: 81pp; English.
XX
XX This sequence is Fas ligand (FasL) mutant delta 2. This protein was made
XX by deleting amino acids 126-135 (inclusive) of the wild-type FasL
XX sequence AAV28594. The FasL protein activates the Fas receptor which
XX induces apoptosis. The FasL deletion mutants (AAV28595-Y28597) are
XX non-cleavable. The mutations inhibit the proteolytic cleavage of FasL
XX and have the capacity to activate a Fas receptor-mediated pathway.
XX Claimed mutants of FasL or an organ or tissue expressing a mutant can be
XX used to alleviate symptoms of a disorder characterized by inadequate or
XX inappropriate stimulation of a Fas receptor-mediated pathway, such as
XX apoptosis, in a tissue or organ. In particular, FasL mutants can be used
XX to treat intolerance to a graft in a patient. The FasL mutants protect
XX an organ or tissue from autoimmune destruction.
XX Note: The present sequence is not shown in the specification but it
XX has been derived from the wild-type FasL sequence AAV28594 shown in the
XX specification.
SQ
Query Match 100.0%; Score 50; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KEIAELRESTS 11
DB 109 KEIAELRESTS 119
RESULT 13
AAV28595
ID AAV28595 standard; Protein; 277 AA.
XX
XX AAV28595;
XX
XX 26-NOV-1999 (first entry)
XX
XX Fas ligand (FasL) mutant delta 1.
XX
XX Fas ligand; FasL; apoptosis; non-cleavable; graft intolerance;
XX autoimmune destruction; cleavage site; mutein.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 125..126 /note="Deletion of 4 wild-type amino acids (126-129)"
XX
XX WO9936079-A1.
XX
XX 22-JUL-1999.
XX
XX 07-JAN-1999; 99WO-US00667.
XX
XX 14-JAN-1998; 98US-0006755.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Baekkeskov S, Braat D, Kang S, Stock PG;
XX
XX WPI; 1999-468942/39.
XX
XX A non-cleavable Fas ligand polypeptide that has capacity to activate
XX Fas receptor-mediated apoptosis
XX Example 1; Page -: 81pp; English.
XX
XX This sequence is Fas ligand (FasL) mutant delta 1. This protein was made
XX by deleting amino acids 126-129 (inclusive) of the wild-type FasL
XX sequence AAV28594. The FasL protein activates the Fas receptor which
XX induces apoptosis. The FasL deletion mutants (AAV28595-Y28597) are
XX

```

CC non-cleavable. The mutations inhibit the proteolytic cleavage of FasL  
 CC and have the capacity to activate a Fas receptor-mediated pathway.  
 CC Claimed mutants of FasL or an organ or tissue expressing a mutant can be  
 CC used to alleviate symptoms of a disorder characterized by inadequate or  
 CC inappropriate stimulation of a Fas receptor-mediated pathway, such as  
 CC apoptosis, in a tissue or organ. In particular, FasL mutants can be used  
 CC to treat intolerance to a graft in a patient. The FasL mutants protect  
 CC an organ or tissue from autoimmune destruction.  
 CC Note: The present sequence is not shown in the specification but it  
 CC has been derived from the wild-type FasL sequence AAY8594 shown in the  
 CC specification.

XX Sequence 277 AA:

Query Match 100.0%; Score 50; DB 20; Length 277;

Best Local Similarity 100.0%; Pred. No. 0.088;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
 |||

DB 109 KELAELRESTS 119

RESULT 14

AAV04372 standard; Protein; 277 AA.

XX AAV04372;

XX 24-JUN-1999 (first entry)

XX Human Fas ligand derivative #2.

XX Human: Fas ligand derivative; antagonist; apoptosis modulator;

XX apoptosis inducer; cytotoxic activity; protease resistant; cancer;

XX viral infection; autoimmune disease; gene therapy.

XX Homo sapiens.

XX Synthetic.

XX WO9914325-A1.

XX 25-MAR-1999.

XX 17-SEP-1998; 98WO-JP04187.

XX 17-SEP-1997; 97JP-0252541.

XX (MOCH ) MOCHIDA PHARM CO LTD.

XX (OSAB-) OSAKA BIOSCIENCE INST.

XX Nagata S, Tanaka M;

XX WPI, 1999-229531/19.

XX N-PSDB; AAX33117.

XX Protease-resistant Fas ligand derivatives used for prevention of,

XX e.g. cancer

XX Claim 4; Page 48-49; 60pp; Japanese.

XX The present sequence is a protease-resistant Fas ligand derivative

XX in which a region of human Fas ligand which is susceptible to protease

XX attack has been deleted. The present invention also describes apoptosis

XX modulators containing soluble Fas ligand. The modification in the Fas

XX ligand renders it resistant to the action of proteases such as the

XX metalloproteinase which in vivo cleaves the active membrane-bound Fas

XX ligand (which is active as an apoptosis inducer). The Fas ligand can be

XX used for the prevention and treatment of diseases such as cancer, viral

XX infection and autoimmune disease, e.g. by introduction of DNA encoding

XX the modified Fas ligand into effector cells using a suitable gene

XX therapy vector.

XX Sequence 277 AA:

Query Match 100.0%; Score 50; DB 20; Length 277;

Best Local Similarity 100.0%; Pred. No. 0.088;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
 |||

DB 109 KELAELRESTS 119

RESULT 15

AAV77281 standard; Protein; 281 AA.

XX AAV77281;

XX 05-DEC-1995 (first entry)

XX Human Fas-L protein.

XX Fas ligand; Fas-L; cell surface protein; autoimmune disease;

XX self-tolerance.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

XX 1..80

XX /Label= Cytoplasmic\_domain

XX 81..105

XX /Label= Transmembrane\_domain

XX 106..281

XX /Label= Extracellular\_domain

XX WO9518819-A.

XX 13-JUL-1995.

XX 06-JAN-1995; 95WO-US00362.

XX 01-FEB-1994; 94US-0190559.

XX 07-JAN-1994; 94US-0179138.

XX (IMM) IMMUNEX CORP.

XX Goodwin RG;

XX WPI, 1995-255032/33.

XX N-PSDB; AAQ91311.

XX Human and murine DNA encoding ligand(s) binding to cell surface protein

XX Fas - useful for studying auto-immune disorder(s) and development of

XX self-tolerance.

XX Claim 3; Page 25-27; 38pp; English.

XX A cDNA library prepd. from human peripheral blood lymphocyte mRNA was

XX screened with a murine Fas-L probe. Isolated clones were amplified

XX by PCR, and a desired fragment ligated into vector pBluescript SK.

XX The cDNA insert of recombinant vector Fas-L/pBS, deposited as

XX E. coli ATCC 69527, is given in AAQ91311. Recombinant Fas-L was

XX expressed in CV1-EBNA-1 cells.

XX Sequence 281 AA:

XX Query Match 100.0%; Score 50; DB 16; Length 281;

XX Best Local Similarity 100.0%; Pred. No. 0.089;

XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX OY 1 KELAELRESTS 11  
 |||

XX DB 109 KELAELRESTS 119

Fri Mar 14 10:16:27 2003

us-10-050-200-3.rag

Page 9

Search completed: March 13, 2003, 18:49:12  
Job time : 30.2 secs

---

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:40:08 ; Search time 26.5455 Seconds  
(without alignments)  
50.197 Million cell updates/sec

Title: US-10-050-200-6

Perfect score: 45

Sequence: 1-EKARVLAEM 10

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

A\_Geneseq.101002.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	88.9	9	22	ABP12751
2	40	88.9	9	22	ABP17961
3	40	88.9	10	23	ABR80017
4	40	88.9	10	23	AA19118
5	40	88.9	10	23	AA175795
6	40	88.9	11	22	ABP17190
7	40	88.9	11	23	AAU80096
8	40	88.9	11	23	ABR76233
9	40	88.9	12	18	AAW36881
10	40	88.9	12	22	ABP17219

11	40	88.9	12	22	ABP17220	HIV B27 super moti
12	40	88.9	14	13	AA25633	HIV1, gag 357-370.
13	40	88.9	15	16	AA68705	T cell epitope, P2
14	40	88.9	15	18	AAW25875	HIV T-helper deter
15	40	88.9	15	19	AAW67391	HIV-1 peptide epit
16	40	88.9	15	20	AA39674	HIV1 gag protein e
17	40	88.9	15	20	AAW98894	HIV-1 vaccine synt
18	40	88.9	15	22	ABP24909	HIV DR 3a motif ga
19	40	88.9	15	22	ABP24915	HIV DR 3a motif ga
20	40	88.9	25	19	AAW82529	HIV-1 p24 epitope
21	40	88.9	30	23	AAU84422	HIV gag segment 24
22	40	88.9	32	16	AAW68566	P24M/V3 consensus
23	40	88.9	32	18	AAW25846	Chimaeric T/B cell
24	40	88.9	32	19	AAW67362	HIV-1 peptide epit
25	40	88.9	32	20	AA39767	HIV1 chimeric pept
26	40	88.9	32	20	AAW99970	HIV-1 vaccine synt
27	40	88.9	34	16	AA68706	P24M/CLTB-56 chime
28	40	88.9	34	16	AA68707	CLTB-56/P24M chime
29	40	88.9	34	16	AA77726	P24M/V3 consensus
30	40	88.9	34	18	AAW25877	Chimaeric T/B cell
31	40	88.9	34	18	AAW25876	Chimaeric T/B cell
32	40	88.9	34	19	AAW67392	HIV-1 peptide epit
33	40	88.9	34	19	AAW67393	HIV-1 peptide epit
34	40	88.9	34	20	AA39722	HIV1 chimeric pept
35	40	88.9	34	20	AA39723	HIV1 chimeric pept
36	40	88.9	34	20	AAW98855	HIV-1 vaccine synt
37	40	88.9	34	20	AAW99903	HIV-1 vaccine synt
38	40	88.9	36	16	AAW77756	CLTB-56/P24M chime
39	40	88.9	36	16	AAW77757	CLTB-56/P24M chime
40	40	88.9	40	17	AAW03358	Human immunodeficie
41	40	88.9	69	12	AA14009	HIV-1 hxb2 gag 318
42	40	88.9	69	12	AA14439	HIV-1 hxb2 gag 318
43	40	88.9	97	22	AA67383	P24c DNA. Human 1
44	40	88.9	100	17	AAW03357	Human immunodeficie
45	40	88.9	172	14	AA43240	Gag sequence. Hum

#### ALIGNMENTS

RESULT 1  
ID ABP12751 standard; Peptide: 9 AA.  
XX  
AC ABP12751:  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE HIV A02 super motif gag peptide #222.  
XX  
XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
KW vpr; vif; tat; cytoxic T lymphocyte; CTL; immune response; epitope;  
KW antigen; vaccine; HIV infection; immunisation; virucide.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
PN WO200124810-A1.  
XX  
PD 12-APR-2001.  
XX  
PE 05-OCT-2000; 2000WO-US27766.  
XX  
PR 05-OCT-1999; 99US-0412863.  
XX  
PA (EPIM-) EPIMMUNE INC.  
XX  
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
PI Baker DM, Cells E, Kubo RT, Grey HM;  
XX  
DR WPI; 2001-354887/37.  
XX  
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
peptide groups, useful for vaccinating against HIV-1 -

XX Claim 32; Page 130; 448pp; English.  
 XX  
 CC The present invention describes a composition (I) comprising a prepared  
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)  
 CC may be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines.  
 CC An additional advantage of an group-based vaccine approach is the ability  
 CC to combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABP1501 to ABP25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention.  
 CC  
 SQ Sequence 9 AA;  
 XX  
 Query Match 88.9%; Score 40; DB 22; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 KARVLAEAM 10  
 |||||  
 DB 1 KARVLAEAM 9  
 RESULT 2  
 ABB17961  
 ID ABB17961 standard; Peptide; 9 AA.  
 XX  
 AC ABB17961;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE HIV B58 super motif gag peptide #180.  
 XX  
 KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
 KW vpr; vif; tat; cytoxic T lymphocyte; CTL; Immune response; epitope;  
 KW antigen; vaccine; HIV infection; Immunisation; virucide.  
 OS  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO200124810-A1.  
 PD 12-APR-2001.  
 XX  
 PF 05-OCT-2000; 2000WO-US27766.  
 XX  
 PR 05-OCT-1999; 99US-0412863.  
 XX  
 PA (EPTM-) EPTMUNE INC.  
 XX  
 PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Cells E, Kubo RT, Grey HM;  
 XX  
 DR WPI; 2001-354887/37.  
 XX  
 PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
 PT peptide groups, useful for vaccinating against HIV-1.  
 XX  
 PS Claim 32; Page 237; 448pp; English.  
 XX

CC The present invention describes a composition (I) comprising a prepared  
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)  
 CC may be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines.  
 CC An additional advantage of an group-based vaccine approach is the ability  
 CC to combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABP1501 to ABP25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention.  
 CC  
 SQ Sequence 9 AA;  
 XX  
 Query Match 88.9%; Score 40; DB 22; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 KARVLAEAM 10  
 |||||  
 DB 1 KARVLAEAM 9  
 RESULT 3  
 ABB80017  
 ID ABB80017 standard; peptide; 10 AA.  
 XX  
 AC ABB80017;  
 XX  
 DT 17-JUL-2002 (first entry)  
 XX  
 DE Cathepsin S inhibition assay substrate peptide.  
 XX  
 KW Cathepsin S; anti-allergic; antiasthmatic; dermatological; pyrazole;  
 KW atopic allergy; hay fever; asthma; atopic dermatitis; food allergy;  
 KW allergy; dust; pollen; mold; pet dander; pet hair.  
 XX  
 OS Unidentified.  
 OS  
 FH  
 FT Key Location/Qualifiers  
 FT Modified-site 10  
 FT /note="C-terminal amide"  
 XX  
 PN WO200220011-A2.  
 PD 14-MAR-2002.  
 XX  
 PF 05-SEP-2001; 2001WO-US27429.  
 XX  
 PR 06-SEP-2000; 2000US-230407P.  
 PR 10-AUG-2001; 2001US-0927324.  
 XX  
 PA (ORTH ) ORTHO-MCNEIL PHARM INC.  
 XX  
 PI Butler CR, Cai H, Edwards JP, Grice CA, Gu Y, Gustin DJ;  
 PI Karlsson U, Khattaya H, Medina SP, Plo BA, Sehon CA, Sun S;  
 PI Tays KL, Thurmond RL, Wei J;  
 XX  
 DR WPI; 2002-393814/42.  
 XX  
 PT Treatment of allergic conditions e.g. atopic dermatitis or asthma  
 PT comprises administration of pyrazole derivative that inhibits cathepsin

PT S -  
 XX  
 PS Example 54: Page 134; 165pp; English.  
 CC The invention relates to the treatment of allergic conditions, comprising  
 CC the administration of a pyrazole derivative. The activity of derivatives  
 CC of the invention may be described as, antiallergic, antiasthmatic and  
 CC dermatological. They act as cathepsin S inhibitors and may be used for  
 CC the treatment of allergic conditions including atopic allergies, e.g. hay  
 CC fever, asthma, atopic dermatitis, food allergies and allergies to dust,  
 CC pollen, mold, pet dander or pet hair. The current sequence represents a  
 CC cathepsin S inhibition assay substrate peptide.  
 SO Sequence 10 AA;  
 Query Match 88.9%; Score 40; DB 23; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EKARVLAEA 9  
 DB 1 EKARVLAEA 9  
 RESULT 4  
 AAE19118  
 ID AAE19118 standard; peptide; 10 AA.  
 AC AAE19118;  
 XX  
 DT 21-MAY-2002 (first entry)  
 DE Protease peptide.  
 XX  
 KM Luciferase; enzyme; protease; cell growth; apoptosis; therapeutic;  
 KM cell death.  
 XX  
 OS Unidentified.  
 PN WO200206458-A2.  
 PD 24-JAN-2002.  
 XX  
 PF 17-JUL-2001; 2001MO-US22478.  
 PR 19-JUL-2000; 2000US-0619047.  
 PA (CHEM-) CHEMICON INT INC.  
 PI Leng J;  
 DR WPI; 2002-195809/25.  
 PT New purified luciferase polypeptide, useful for identifying protease  
 PT activity modulators, comprises recognition site cleavable by protease,  
 PT where cleavage of polypeptide results in decreased luciferase activity  
 PT  
 PS Claim 63; Page 24; 56pp; English.  
 CC The invention relates to a purified polypeptide having luciferase  
 CC activity and a recognition site specifically cleavable by a protease,  
 CC where cleavage results in a decrease in luciferase activity. The  
 CC polypeptide comprises a localisation sequence which is linked to the  
 CC luciferase polypeptide by the cleavable recognition sequence. The  
 CC polypeptide is useful for identifying a protease (Caspase-3) activity  
 CC modulator, an inhibitor of apoptosis and for detecting luciferase  
 CC activity in a sample. The polypeptide is used for characterising and  
 CC identifying cellular processes associated with metabolism, cell growth  
 CC and cell death e.g. apoptosis and for measuring protease activity. The  
 CC protease measurement methods are useful for characterising, identifying  
 CC cellular biochemical pathways as well as identifying diagnostic and  
 CC agents for modulating diseases or disorders associated with

CC biochemical pathways. The polypeptide is also useful as a substrate to  
 CC study agents or conditions that cleave the recognition site and for  
 CC determining amount of protease in a sample and for monitoring the  
 CC activity of protease inside a cell that expresses a recombinant  
 CC luciferase. The present sequence is a protease peptide used in the  
 CC invention.  
 SO Sequence 10 AA;  
 Query Match 88.9%; Score 40; DB 23; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KARVLAEM 10  
 DB 1 KARVLAEM 9  
 RESULT 5  
 AAU75795  
 ID AAU75795 standard; Peptide; 10 AA.  
 AC AAU75795;  
 XX  
 DT 08-MAY-2002 (first entry)  
 DE Human immunodeficiency virus Gag protein protease cleavage sequence #2.  
 XX  
 KM Gag protein; antiviral; nucleic acid construct; viral protease;  
 KM protease recognition site; PRS; viral drug resistance; HIV;  
 KM human immunodeficiency virus.  
 OS Human immunodeficiency virus type 1.  
 FH Key Location/Qualifiers  
 FT Cleavage-site 5  
 FT /label= HIV-protease\_cleavage-site  
 XX  
 PN WO200210430-A2.  
 PD 07-FEB-2002.  
 XX  
 PF 30-JUL-2001; 2001MO-IL00702.  
 PR 31-JUL-2000; 2000US-0629969.  
 PA (AMID-) AMIDUT LTD.  
 PI Appel E;  
 DR WPI; 2002-188753/24.  
 PT New nucleic acid construct for detecting anti-viral drugs, comprises a  
 PT polynucleotide cassette encoding a chimeric polypeptide with two  
 PT polypeptide sequences and a protease recognition site cleavable by a  
 PT virally encoded protease  
 PT  
 PS Example 1; Page 32; 93pp; English.  
 CC The present invention relates to a new nucleic acid construct that  
 CC comprises a polynucleotide cassette encoding a chimeric polypeptide.  
 CC The chimeric polypeptide comprises a first polypeptide sequence, a  
 CC second polypeptide sequence translationally fused to the first sequence,  
 CC and a protease recognition site (PRS) cleavable by a virally encoded  
 CC protease, where cleavage of the PRS leads to a detectable signal. The  
 CC nucleic acid construct of the invention is useful for uncovering  
 CC molecules having antiviral activity or for determining viral drug  
 CC resistance and the chimeric peptide is also useful for detecting the  
 CC presence of a virus in a cell. Other uses of the invention are detecting  
 CC viral encoded protease found in infected cells or detecting the presence  
 CC or absence of viral infection. The nucleic acid construct is useful for  
 CC phenotypic testing of human immunodeficiency virus (HIV) drug resistance.  
 CC The invention enables screening of molecules in an easy and rapid manner

CC and allows efficient detection of the presence of viral protease and  
 CC viral particles within cells. The construct is specific, sensitive and  
 CC lacks background enzymatic activity in the absence of human  
 CC immunodeficiency virus protease, and is suitable for detection of  
 CC specific viral strain isolates even under low viral load conditions. When  
 CC used for phenotypic testing of HIV drug resistance it delivers accurate  
 CC results within 24 hours. The present amino acid sequence represents one  
 CC of several (AAU75794-AAU75798) HIV gag protein protease cleavage  
 CC sequences that were used in the invention for HIV detection.

CC  
 CC Sequence 10 AA;  
 CC  
 CC Query Match 88.9%; Score 40; DB 23; Length 10;  
 CC Best Local Similarity 100.0%; Pred. No. 0.13;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 DB 1 KARVLAEM 9  
 |||||

RESULT 6  
 ABP17190  
 ID ABP17190 standard; Peptide: 11 AA.  
 AC ABP17190;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE HIV B27 super motif gag peptide #66.  
 XX  
 KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;  
 KW vpr; vif; tat; cytototoxic T lymphocyte; CTL; immune response; epitope;  
 KW antigen; vaccine; HIV infection; immunisation; vifucide.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO200124810-A1.  
 PD 12-APR-2001.  
 XX  
 PF 05-OCT-2000; 2000WO-US27766.  
 XX  
 PR 05-OCT-1999; 99US-0412863.  
 XX  
 PA (EPI-M-) EPI-MUNE INC.  
 XX  
 PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Cells E, Kubo RT, Grey HM;  
 XX  
 DR WPI; 2001-354887/37.  
 XX  
 PT Vaccines comprising human immunodeficiency virus-1 (HIV-1)  
 PT peptide groups, useful for vaccinating against HIV-1 -  
 XX  
 PS Claim 32; Page 221; 448pp; English.

CC The present invention describes a composition (I) comprising a prepared  
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)  
 CC may be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines.  
 CC An additional advantage of an group-based vaccine approach is the ability  
 CC to combine selected groups (CTL and HTL), and further, to modify the

CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABP1501 to ABP25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention.

CC  
 CC Sequence 11 AA;  
 CC  
 CC Query Match 88.9%; Score 40; DB 22; Length 11;  
 CC Best Local Similarity 100.0%; Pred. No. 0.15;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 DB 3 KARVLAEM 11  
 |||||

RESULT 7  
 AAU80096  
 ID AAU80096 standard; peptide: 11 AA.  
 AC AAU80096;  
 XX  
 DT 24-SEP-2002 (first entry)  
 XX  
 DE Human cathepsin S substrate peptide.  
 XX  
 KW Cathepsin S; allergy; pyrazole; hay fever; asthma; atopic dermatitis;  
 KW food allergy.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200220013-A2.  
 PD 14-MAR-2002.  
 XX  
 PF 05-SEP-2001; 2001WO-US27480.  
 XX  
 PR 06-SEP-2000; 2000US-230407P.  
 PR 10-AUG-2001; 2001US-0927188.  
 XX  
 PA (ORTH ) ORTHO-MCNEIL PHARM INC.  
 XX  
 PI Cai H, Edwards JP, Gu Y, Karlsson L, Medina SP, Pio BA, Sun S;  
 PI Thuimond RL, Wei J;  
 XX  
 DR WPI; 2002-527344/56.  
 XX  
 PT Treating allergic conditions, e.g. hay fever, asthma, atopic dermatitis  
 PT and food allergies, comprises administration of a substituted pyrazole  
 XX  
 PS Example 37; Page 92; 115pp; English.

CC The invention relates to a method of treating allergic conditions  
 CC comprising administration of a composition comprising a substituted  
 CC pyrazole (I) or a salt, amide or ester. The method is used for treating  
 CC allergic conditions, e.g. hay fever, asthma, atopic dermatitis and food  
 CC allergies. The present sequence represents human cathepsin S  
 CC substrate peptide used in an assay to demonstrate the method of the  
 CC invention.

CC  
 CC Sequence 11 AA;  
 CC

Query Match 88.9%; Score 40; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKARVLAEA 9  
 |||||  
 DB 1 EKARVLAEA 9

# RESULT 8

ABB76233  
 ID ABB76233 standard; Peptide: 11 AA.

AC ABB76233;

DT 09-AUG-2002 (first entry)

DE Substrate for cathepsin S inhibition assay.

KM Cathepsin S; inhibitor; allergy; hay fever; asthma; food allergy;  
 KM atopic dermatitis; antiallergic; antisthmatic; dermatological;  
 KW antiinflammatory; immunosuppressive.

OS Synthetic.

PH Key Location/Qualifiers

FT Modified-site 1 /note="adens-glutamic acid"

FT Modified-site 11 /note="Dabcy1 lysine-amide"

PN WO200220002-A2.

PD 14-MAR-2002.

PF 05-SEP-2001; 2001WO-US27441.

PR 06-SEP-2000; 2000US-230407P.

PA (ORTH ) ORTHO-MCNEIL PHARM INC.

PI Gu Y, Karlsson L, Sun S, Thurmond R;

DR WPI; 2002-443877/47.

PT Use of cathepsin S inhibitor for treating an allergic condition, e.g.  
 hay fever, asthma, atopic dermatitis or a food allergy

PS Example 1; Page 29; 53pp; English.

XX The present sequence is a peptide used as substrate for a  
 CC cathepsin S inhibition assay. The Km for the substrate is around  
 CC 5 uM but the presence of substrate inhibition makes kinetic  
 CC analysis difficult. With 20 uM substrate, the assay rate is  
 CC linear over the range of 1-8 ng cathepsin S in 100 uL reaction.  
 CC Using 2 ng/well of cathepsin S, the production of product is  
 CC linear and yields about 7-fold signal after 20 min with only 20%  
 CC loss of substrate. Primary assays are run by quenching the  
 CC reaction after 20 min and then measuring the fluorescence. For  
 CC other assays, measurements are taken every min for 20 min. The  
 CC rate is calculated from the slope of the increase and the percent  
 CC inhibition is calculated from this. The present invention relates  
 CC to the use of cathepsin S inhibitors to treat an allergic  
 CC condition, especially an atopic IGE-mediated allergic condition  
 CC such as hay fever, asthma, atopic dermatitis or a food allergy  
 CC (claimed). The allergens may include dust, pollen, mould, and pet  
 CC dander or pet hair.

XX Sequence 11 AA;

Query Match 88.9%; Score 40; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.15; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

OY 1 EKARVLAEA 9  
 |||||  
 DB 1 EKARVLAEA 9

# RESULT 9

AAW36881  
 ID AAW36881 standard; Peptide: 12 AA.

AC AAW36881;

DT 11-MAY-1998 (first entry)

DE HIV protease cleavage recognition site.

KM Ricin; toxin; antiviral; virucide; retrovirus; protease; HIV;  
 KM AIDS; infection; therapy; linker.

OS Synthetic.

PN WO9741233-A1.

PD 06-NOV-1997.

PF 29-APR-1997; 97WO-CA00288.

PR 30-APR-1996; 96US-0016509.

PA (CANG-) CANGENE CORP.

PI Borgford T;

DR WPI; 1997-549735/50.

PT DNAs encoding ricin like toxins A and B - are linked via linker  
 containing cleavage site for retroviral protease, used to inhibit or  
 PT destroy mammalian cells infected with retrovirus

PS Claim 5; Page 41; 105pp; English.

XX This claimed peptide is a cleavage recognition site for a HIV  
 CC protease. It is utilised as a linker between the A and B chains  
 CC of a ricin-like protein in a novel recombinant protein. A nucleic  
 CC acid (see AAW97900) encoding such a construct is obtained by PCR  
 CC mutagenesis of the wild-type ricin linker sequence. The invention  
 CC provides novel recombinant proteins which incorporate the A and B  
 CC chains of a ricin-like toxin (preferably the A and B chains of  
 CC ricin) linked by a heterologous linker sequence containing a  
 CC cleavage recognition site for a retroviral protease such as HIV  
 CC protease (see AAW36880-82), HTLV-I (see AAW36883-43) or HTLV-II (see  
 CC AAW36885-65). The recombinant proteins selectively inhibit or  
 CC destroy mammalian cells infected with a retrovirus such as cancer  
 CC cells associated with HTLV or cells associated with HIV. The  
 CC recombinant proteins are non-toxic until the ricin A chain is  
 CC liberated from the B chain by a retroviral protease, and thus can  
 CC be used to specifically target infected cells without the need  
 CC for a cell binding component.

XX Sequence 12 AA;

Query Match 88.9%; Score 40; DB 18; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.16;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 |||||

DB 2 KARVLAEM 10

# RESULT 10

ABP17219  
 ID ABP17219 standard; Peptide: 12 AA.

```

XX AC ABP17219;
XX
XX 15-JUL-2002 (first entry)
XX
XX DE HIV B27 super motif gag peptide #95.
XX
XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;
XX vpr; vif; tat; cytoxic T lymphocyte; CTL; immune response; epitope;
XX antigen; vaccine; HIV infection; immunisation; virucide.
XX
XX Human immunodeficiency virus type 1.
XX
XX WO200124810-A1.
XX
XX 12-APR-2001.
XX
XX 05-OCT-2000; 2000WO-US27766.
XX
XX 05-OCT-1999; 99US-0412863.
XX
XX (EPIM-) EPIMUNE INC.
XX
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Cells E, Kubo RT, Grey HM;
XX WPI; 2001-354887/37.
XX
XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
XX peptide groups, useful for vaccinating against HIV-1.
XX
XX Claim 32: Page 221; 448pp; English.
XX
XX The present invention describes a composition (I) comprising a prepared
XX human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
XX sequence selected from 51 defined amino acid sequences (ABP25347 to
XX ABP25397). (I) has virucide activity and can be used in vaccines. (I)
XX may be used for immunising subjects against HIV-1 infections. The use of
XX group-based vaccines has several advantages over traditional vaccines,
XX particularly when compared to the use of whole antigens in vaccine
XX compositions. There is evidence that the immune response to whole
XX antigens is directed largely toward variable regions of the antigen,
XX allowing for immune escape due to mutations. The groups for inclusion in
XX an group-based vaccine may be selected from conserved regions of viral or
XX tumour-associated antigens, which therefore reduces the likelihood of
XX escape mutants. Furthermore, immunosuppressive groups that may be present
XX in whole antigens can be avoided with the use of group-based vaccines.
XX An additional advantage of an group-based vaccine approach is the ability
XX to combine selected groups (CTL and RTL), and further, to modify the
XX composition of the groups, achieving, for example, enhanced
XX immunogenicity. Accordingly, the immune response can be modulated, as
XX appropriate, for the target disease. Similar engineering of the response
XX is not possible with traditional approaches. ABP1501 to ABP25412
XX represent peptide sequences used in the exemplification of the present
XX invention.
XX
XX SQ Sequence 12 AA;
XX
XX Query Match 88.9%; Score 40; DB 22; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 0.16;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 KARVLAEM 10
   |||||
DB 4 KARVLAEM 12

```

```

RESULT 11
ABP17220
ID ABP17220 standard; Peptide; 12 AA.
XX
XX AC ABP17220;
XX

```

```

DT 15-JUL-2002 (first entry)
XX
XX DE HIV B27 super motif gag peptide #96.
XX
XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;
XX vpr; vif; tat; cytoxic T lymphocyte; CTL; immune response; epitope;
XX antigen; vaccine; HIV infection; immunisation; virucide.
XX
XX Human immunodeficiency virus type 1.
XX
XX WO200124810-A1.
XX
XX 12-APR-2001.
XX
XX 05-OCT-2000; 2000WO-US27766.
XX
XX 05-OCT-1999; 99US-0412863.
XX
XX (EPIM-) EPIMUNE INC.
XX
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Cells E, Kubo RT, Grey HM;
XX WPI; 2001-354887/37.
XX
XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
XX peptide groups, useful for vaccinating against HIV-1.
XX
XX Claim 32: Page 221; 448pp; English.
XX
XX The present invention describes a composition (I) comprising a prepared
XX human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
XX sequence selected from 51 defined amino acid sequences (ABP25347 to
XX ABP25397). (I) has virucide activity and can be used in vaccines. (I)
XX may be used for immunising subjects against HIV-1 infections. The use of
XX group-based vaccines has several advantages over traditional vaccines,
XX particularly when compared to the use of whole antigens in vaccine
XX compositions. There is evidence that the immune response to whole
XX antigens is directed largely toward variable regions of the antigen,
XX allowing for immune escape due to mutations. The groups for inclusion in
XX an group-based vaccine may be selected from conserved regions of viral or
XX tumour-associated antigens, which therefore reduces the likelihood of
XX escape mutants. Furthermore, immunosuppressive groups that may be present
XX in whole antigens can be avoided with the use of group-based vaccines.
XX An additional advantage of an group-based vaccine approach is the ability
XX to combine selected groups (CTL and RTL), and further, to modify the
XX composition of the groups, achieving, for example, enhanced
XX immunogenicity. Accordingly, the immune response can be modulated, as
XX appropriate, for the target disease. Similar engineering of the response
XX is not possible with traditional approaches. ABP1501 to ABP25412
XX represent peptide sequences used in the exemplification of the present
XX invention.
XX
XX SQ Sequence 12 AA;
XX
XX Query Match 88.9%; Score 40; DB 22; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 0.16;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 KARVLAEM 10
   |||||
DB 4 KARVLAEM 12

```

```

RESULT 12
AAR25633
ID AAR25633 standard; Protein; 14 AA.
XX
XX AC AAR25633;
XX
XX DT 17-DEC-2001 (updated)
XX DT 19-JAN-1993 (first entry)
XX

```

```
DE HIV1, gag 357-370.
XX
KM Polar; hydrophobic; hydrolysis; inhibitor; therapeutic.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Cleavage-site 7..8
XX
PN USN771554-N.
XX
PD 15-APR-1992.
XX
PF 08-OCT-1991; 91US-0771554.
XX
PR 08-OCT-1991; 91US-0771554.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Knight M, Mukherjee AB;
XX
DR WPI; 1992-259289/31.
XX
PT Assay for retroviral aspartyl proteinase - provides rapid
PT accurate and inexpensive method for detecting aspartyl proteinase
PT inhibitors, e.g. for AIDS treatment
XX
PS Disclosure: Page 10; 27pp; English.
XX
CC The sequences given in AAR25632-38 are peptides which are hydrolysed
CC by aspartyl proteinase (AP) in the assay of the invention. These
CC peptides have labelled C-terminal portions which are polar. These
CC portions are separated from the unlabelled hydrophobic N-terminal
CC portions by hydrolysis by AP. Following hydrolysis of these peptides
CC with AP, the hydrolysed polar portion is separated from the unhydrolysed
CC polypeptide and the amount of hydrolysed polypeptide determined. This
CC assay can be used to detect inhibitors of AP with potential as
CC therapeutic agents, eg. for treating AIDS. This assay has a higher
CC sensitivity than previous methods.
CC
CC (Note: Revised entry submitted to correct the patent number format of
CC US Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis-us.html.)
XX
SQ Sequence 14 AA:
Query Match: 88.9%; Score 40; DB 13; Length 14;
Best Local Similarity 100.0%; Pred. NO. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 2 KARVLAEAM 10
| | | | |
Db 3 KARVLAEAM 11
RESULT 13
AAR68705
ID AAR68705 standard; peptide; 15 AA.
XX
AC AAR68705;
DT 07-SEP-1995 (first entry)
XX
DE T cell epitope, P24M.
XX
KW T-cell; epitope; HIV-1; core protein; p24E; B-cell; antigen;
KW gp160; gag; pol; vaccine; multimeric peptide; AIDS; 3D organisation.
XX
OS Human immunodeficiency virus type 1.
XX
MN WO9429339-A.
XX
PD 22-DEC-1994.
```

Pf	08-JUN-1994;	94WO-CA00317.
Pf	09-JUN-1993;	93US-0073378.
Pa	(CONN-) CONNAUGHT LAB LTD.	
Pl	Chong P, Klein MH, Sia CDY;	
Df	WPI; 1995-036400/05.	
Pt	Novel tandem synthetic HIV-1 peptide(s) - comprising T-cell	
Pt	epitope of gag protein linked to B-cell epitope of V3 loop	
Pt	protein of an HIV-1 isolate	
Pf	Claim 4; Page 45; 69pp; English.	
Cc	This sequence represents the T-cell epitope, p24M, which may be linked	
Cc	to a B-cell epitope from the V3 (MN) loop from HIV-1. These chimeric	
Cc	peptides may then be used in the production of HIV-1 vaccines. These	
Cc	peptide sequences may also be used in the production of multimeric	
Cc	peptides in which the linear peptides are C-terminally modified by	
Cc	the addition of a lys residue which is modified on its epsilon amino	
Cc	acid to carry an additional copy of the peptide molecule. The linear	
Cc	and multimeric peptides may be used for the treatment of AIDS by	
Cc	acting to displace the binding of HIV virus to human or animal cells	
Cc	or by disturbing the 3D organisation of the virus.	
Sd	Sequence 15 AA:	
Qy	Query Match	88.9%; Score 40; DB 16; Length 15;
Qy	Best Local Similarity	100.0%; Pred. No. 0.2;
Matches	9; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Db	2 KARVLAEAM 10           3 KARVLAEAM 11	
RESULT 14		
ID	AAW25875	
AC	AAW25875 standard; peptide: 15 AA.	
DT	22-OCT-1997 (first entry)	
DE	HIV T-helper determinant core peptide p24M.	
KV	HIV; human immunodeficiency virus; gag; T-cell; B-cell; epitope; env;	
KV	V3 loop; vaccine; determinant; chimaeric.	
OS	Synthetic.	
PN	US5639854-A.	
PD	17-JUN-1997.	
Pe	09-JUN-1993; 93US-0073378.	
PR	09-JUN-1994; 94US-0257528.	
PR	09-JUN-1993; 93US-0073378.	
PA	(CONN-) CONNAUGHT LAB LTD.	
Pl	Chong P, Klein MH, Sia CDY;	
Df	WPI; 1997-332082/30.	
Pt	Tandem synthetic HIV peptide(s) useful as immunogens - comprising	
Pt	gag protein T-cell epitope linked to env protein B-cell epitope	
Pf	Claim 1; Column 73; 41pp; English.	

XX The invention relates to new synthetic peptides comprising at least one  
 CC amino acid sequence comprising an HIV gag protein T-cell epitope linked  
 CC at its C- or N-terminus to an amino acid sequence comprising a B-cell  
 CC epitope of the V3 loop of an HIV env protein, which can be used to  
 CC generate vaccines against HIV-1. The T-cell epitope sequence is pref.  
 CC selected from the T-helper determinant core peptides P24E, P24N, P24L,  
 CC P24M and P24H while the B-cell epitopes are derived from HIV strains  
 CC including CTLB-56, V3MN, CTLB-29, CTLB-55, SF2, IAI, IIB, RF, 26, 2054,  
 CC 1714 and BX08. The peptides are chimeric and can be linked to a  
 CC branched lys backbone. This sequence represents the T-helper determinant  
 CC core peptide P24M.

XX  
 SQ Sequence 15 AA;

Query Match 88.9%; Score 40; DB 18; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 |||||  
 Db 3 KARVLAEM 11

RESULT 15

AAW67391  
 ID AAW67391 standard; peptide: 15 AA.

AC AAW67391;

DT 25-JAN-1999 (first entry)

DE HIV-1 peptide epitope p24M.

KM Immunogen; vaccine; HIV-1; T-cell; B-cell; epitope; core protein; gp120;  
 V3 loop.

OS Human immunodeficiency virus type 1.

PN US5817754-A.

PD 06-OCT-1998.

PF 05-JUN-1995; 95US-0464329.

PR 09-JUN-1994; 94US-0257528.

PR 09-JUN-1993; 93US-0073378.

PR 05-JUN-1995; 95US-0464329.

PA (CONN-) CONNAUGHT LAB LTD.

PI Chong P, Klein MH, Sia CDY;

DR WPI: 1998-556461/47.

PT Synthetic human immunodeficiency virus-1 peptide(s) - containing  
 PT T-cell epitope and B-cell epitope(s) are candidate vaccines against  
 PT HIV-1

PS Disclosure; Column 23-24; 40pp; English.

XX The invention relates to a novel immunogenic composition for use in  
 CC vaccines for the treatment of HIV-1 comprising an HIV-1-derived T-cell  
 CC epitope linked to an HIV-1-derived B-cell epitope. The T-cell epitopes  
 CC are generally designed based on the p24 core protein and the B-cell  
 CC epitopes from the V3 loop of the gp120 protein from various HIV-1  
 CC strains. This peptide is a T-cell epitope derived from the HIV-1 strain  
 CC MN p24 core protein.

SQ Sequence 15 AA;

Query Match 88.9%; Score 40; DB 19; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
 |||||  
 Db 3 KARVLAEM 11

Search completed: March 13, 2003, 18:49:14  
 Job time : 26.5455 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:46:52 ; Search time 8.90909 Seconds

(without alignments)  
33.026 Million cell updates/sec

Title: US-10-050-200-6

Perfect score: 45

Sequence: 1 EKARVLAEM 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5a COMB pep:\*  
2: /cgn2\_6/prodata/2/1aa/5b COMB pep:\*  
3: /cgn2\_6/prodata/2/1aa/6a COMB pep:\*  
4: /cgn2\_6/prodata/2/1aa/6b COMB pep:\*  
5: /cgn2\_6/prodata/2/1aa/6c COMB pep:\*  
6: /cgn2\_6/prodata/2/1aa/6d COMB pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	88.9	10	2	US-08-792-553-4
2	40	88.9	10	4	US-09-053-941-15
3	40	88.9	10	4	US-09-817-413-15
4	40	88.9	12	4	US-09-147-208-21
5	40	88.9	15	1	US-08-257-528B-76
6	40	88.9	15	1	US-08-460-602A-76
7	40	88.9	15	1	US-08-463-966A-76
8	40	88.9	15	1	US-08-465-217A-76
9	40	88.9	15	2	US-08-464-329A-76
10	40	88.9	15	2	US-08-462-507A-76
11	40	88.9	15	2	US-08-467-881A-76
12	40	88.9	32	1	US-08-257-528B-47
13	40	88.9	32	1	US-08-460-602A-47
14	40	88.9	32	1	US-08-463-966A-47
15	40	88.9	32	1	US-08-465-217A-47
16	40	88.9	32	2	US-08-464-329A-47
17	40	88.9	32	2	US-08-462-507A-47
18	40	88.9	32	2	US-08-467-881A-47
19	40	88.9	34	1	US-08-257-528B-77
20	40	88.9	34	1	US-08-257-528B-78
21	40	88.9	34	1	US-08-460-602A-77
22	40	88.9	34	1	US-08-460-602A-78
23	40	88.9	34	1	US-08-463-966A-77
24	40	88.9	34	1	US-08-463-966A-78
25	40	88.9	34	1	US-08-465-217A-77
26	40	88.9	34	1	US-08-465-217A-78
27	40	88.9	34	2	US-08-464-329A-77

28	40	88.9	34	2	US-08-464-329A-78	Sequence 78, Appl
29	40	88.9	34	2	US-08-462-507A-77	Sequence 77, Appl
30	40	88.9	34	2	US-08-462-507A-78	Sequence 78, Appl
31	40	88.9	34	2	US-08-467-881A-77	Sequence 77, Appl
32	40	88.9	34	2	US-08-467-881A-78	Sequence 78, Appl
33	40	88.9	97	4	US-09-621-625A-27	Sequence 6, Appl1
34	40	88.9	163	4	US-09-117-217-8	Sequence 8, Appl1
35	40	88.9	172	2	US-08-117-981-6	Sequence 6, Appl1
36	40	88.9	172	2	US-08-477-081-6	Sequence 6, Appl1
37	40	88.9	172	2	PCT-US93-02142-6	Sequence 25, Appl
38	40	88.9	197	4	US-09-621-625A-25	Sequence 4, Appl1
39	40	88.9	275	1	US-08-589-446-4	Sequence 4, Appl1
40	40	88.9	275	1	US-08-444-882-4	Sequence 4, Appl1
41	40	88.9	275	1	US-08-389-459A-4	Sequence 4, Appl1
42	40	88.9	275	3	US-08-987-867A-4	Sequence 4, Appl1
43	40	88.9	292	4	US-09-621-625A-2	Sequence 2, Appl1
44	40	88.9	292	4	US-09-621-625A-17	Sequence 17, Appl1
45	40	88.9	294	4	US-09-370-368-3	Sequence 3, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-792-553-4  
; Sequence 4, Application US/08792553  
; Patent No. 5981200  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; TITLE OF INVENTION: Tandem Fluorescent Protein Constructs  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/792,553  
; FILING DATE: 31-JAN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lisa A. Hallie, Ph.D.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07257/041001/OC 96-160-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-678-5099  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-792-553-4  
Query Match 88.9%; Score 40; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 2 KARVLAEM 10  
DB 1 KARVLAEM 9

RESULT 2  
US-09-053-941-15  
; Sequence 15, Application US/09053941  
; Patent No. 6271354  
; GENERAL INFORMATION:  
; APPLICANT: SRINIVASAN, ALGARSAMY  
; APPLICANT: KOPROWSKI, HILARY  
; TITLE OF INVENTION: CHIMERIC VIRAL PROTEINS  
; FILE REFERENCE: Chimeric Viral Proteins  
; CURRENT APPLICATION NUMBER: US/09/053,941  
; CURRENT FILING DATE: 1998-04-02  
; EARLIER APPLICATION NUMBER: 60/043,380  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus type 1  
; PUBLICATION INFORMATION:  
; JOURNAL: Meth. Enzymol.  
; VOLUME: 38  
; PAGES: 299-  
; DATE: 1974  
US-09-053-941-15

Query Match 88.9%; Score 40; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEAM 10  
DB 1 KARVLAEAM 9

RESULT 3  
US-09-817-413-15  
; Sequence 15, Application US/09817413  
; Patent No. 6436648  
; GENERAL INFORMATION:  
; APPLICANT: SRINIVASAN, ALGARSAMY  
; APPLICANT: KOPROWSKI, HILARY  
; TITLE OF INVENTION: CHIMERIC VIRAL PROTEINS  
; FILE REFERENCE: Chimeric Viral Proteins  
; CURRENT APPLICATION NUMBER: US/09/817,413  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: 60/043,380  
; PRIOR FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus type 1  
; PUBLICATION INFORMATION:  
; JOURNAL: Meth. Enzymol.  
; VOLUME: 38  
; PAGES: 299-  
; DATE: 1974  
US-09-817-413-15

Query Match 88.9%; Score 40; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEAM 10  
DB 1 KARVLAEAM 9

RESULT 4  
US-09-147-208-21  
; Sequence 21, Application US/09147208

Patent No. 6333303  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins  
; NUMBER OF SEQUENCES: 71  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERSKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/147,208  
; FILING DATE: 02-MAR-1999  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rudolph, John R.  
; REGISTRATION NUMBER: 38,003  
; REFERENCE/DOCKET NUMBER: 7841-76  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-147-208-21

Query Match 88.9%; Score 40; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEAM 10  
DB 2 KARVLAEAM 10

RESULT 5  
US-08-257-528B-76  
; Sequence 76, Application US/08257528B  
; Patent No. 5639854  
; GENERAL INFORMATION:  
; APPLICANT: SIA, Charles D.Y.  
; APPLICANT: CHONG, Pele  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/257,528B  
; FILING DATE: 09-JUN-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:

NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-336 MIS:jdb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-257-528B-76

Query Match 88.9%; Score 40; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
Db 3 KARVLAEM 11

RESULT 6  
US-08-460-602A-76  
Sequence 76, Application US/08460602A  
Patent No. 5759769  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & Mcburney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-460,602A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-602A-76  
Query Match 88.9%; Score 40; DB 1; Length 15;

Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
Db 3 KARVLAEM 11

RESULT 7  
US-08-463-966A-76  
Sequence 76, Application US/08463966A  
Patent No. 5795955  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & Mcburney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-463,966A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-463-966A-76

Query Match 88.9%; Score 40; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
Db 3 KARVLAEM 11

RESULT 8  
US-08-465-217A-76  
Sequence 76, Application US/08465217A  
Patent No. 5800822  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele

APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,217A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-486 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ. ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-465-217A-76

Query Match 88.9%; Score 40; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEAM 10  
DB 3 KARVLAEAM 11

RESULT 9  
US-08-464-329A-76  
Sequence 76, Application US/08464329A  
Patent No. 5817754  
GENERAL INFORMATION:  
APPLICANT: SIM, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,329A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-449 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ. ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-464-329A-76

Query Match 88.9%; Score 40; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEAM 10  
DB 3 KARVLAEAM 11

RESULT 10  
US-08-462-507A-76  
Sequence 76, Application US/08462507A  
Patent No. 5876731  
GENERAL INFORMATION:  
APPLICANT: SIM, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,507A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-451 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-462-507A-76

Query Match 88.9%: Score 40; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
Db 3 KARVLAEM 11

RESULT 11  
US-08-467-881A-76  
Sequence 76, Application US/08467881A  
Patent No. 5951986  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,881A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-488 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-467-881A-76  
Query Match 88.9%: Score 40; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
Db 3 KARVLAEM 11

RESULT 12  
US-08-257-528B-47  
Sequence 47, Application US/08257528B  
Patent No. 5639854  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/257,528B  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-336 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-257-528B-47

Query Match 88.9%: Score 40; DB 1; Length 32;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
Db 3 KARVLAEM 11

RESULT 13  
US-08-460-602A-47  
Sequence 47, Application US/08460602A  
Patent No. 5759769  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,602A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-450 MIS:jb  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-602A-47

Query Match 88.9%; Score 40; DB 1; Length 32;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
Db 3 KARVLAEM 11

RESULT 14  
US-08-463-966A-47  
Sequence 47, Application US/08463966A  
Patent No. 5795955  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,966A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378

FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-487 MIS:jb  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-463-966A-47

Query Match 88.9%; Score 40; DB 1; Length 32;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
Db 3 KARVLAEM 11

RESULT 15  
US-08-465-217A-47  
Sequence 47, Application US/08465217A  
Patent No. 5800822  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: CHONG, Pele  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,217A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,528  
FILING DATE: 09-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/073,378  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-486 MIS:jb  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-465-217A-47

Query Match 88.9%; Score 40; DB 1; Length 32;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10

|||||||

Db 3 KARVLAEM 11

Search completed: March 13, 2003, 18:53:52  
 Job time : 8.50909 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:49:22 ; Search time 11.1091 Seconds  
(without alignments)  
53.938 Million cell updates/sec

Title: US-10-050-200-7

Perfect score: 62

Sequence: 1 ERAEQRLKSODL 13

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

199416

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUG\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	62.9	146	10	US-09-764-864-1127
2	38	61.3	235	9	US-10-101-464A-787
3	38	61.3	909	9	US-09-925-299-988
4	37	59.7	909	10	US-09-925-299-988
5	37	59.7	404	9	US-09-902-941-1932
6	37	59.7	464	9	US-09-902-941-1934
7	35	56.5	162	9	US-09-854-133-93
8	35	56.5	162	10	US-09-738-973-93
9	35	56.5	229	9	US-09-986-480-344
10	35	56.5	268	10	US-09-864-761-33475
11	35	56.5	341	10	US-09-916-790-18
12	35	56.5	458	10	US-09-801-574-42
13	34	54.8	529	10	US-09-815-242-11532
14	34	54.8	767	10	US-09-919-497-59
15	34	54.8	831	9	US-09-736-457-1819
16	34	54.8	831	9	US-09-902-941-1819
17	34	54.8	831	9	US-09-849-626-1819
18	34	54.8	872	9	US-09-843-676-8
19	34	54.8	872	9	US-09-843-676-54

20	34	54.8	872	9	US-09-766-253-8	Sequence 8, Appl1
21	34	54.8	872	9	US-09-766-253-54	Sequence 54, Appl1
22	34	54.8	872	9	US-09-438-486-8	Sequence 8, Appl1
23	34	54.8	872	9	US-09-438-486-54	Sequence 54, Appl1
24	34	54.8	872	9	US-10-053-758-8	Sequence 8, Appl1
25	34	54.8	872	9	US-10-053-758-54	Sequence 54, Appl1
26	34	54.8	872	9	US-10-054-295-8	Sequence 8, Appl1
27	34	54.8	872	9	US-10-054-295-54	Sequence 54, Appl1
28	34	54.8	1005	10	US-09-925-301-1335	Sequence 1335, Ap
29	33.5	54.0	1192	10	US-09-815-242-10903	Sequence 10903, A
30	33	53.2	38	9	US-10-099-766-12	Sequence 12, Appl1
31	33	53.2	77	10	US-09-824-178-12	Sequence 12, Appl1
32	33	53.2	117	10	US-09-824-178-3	Sequence 3, Appl1
33	33	53.2	117	12	US-10-016-993-4	Sequence 4, Appl1
34	33	53.2	121	10	US-09-824-178-4	Sequence 4, Appl1
35	33	53.2	121	12	US-10-016-993-3	Sequence 3, Appl1
36	33	53.2	136	10	US-09-864-761-36003	Sequence 36003, A
37	33	53.2	137	10	US-09-864-761-36465	Sequence 36465, A
38	33	53.2	150	10	US-09-925-297-534	Sequence 534, App
39	33	53.2	337	10	US-09-815-242-5617	Sequence 5617, Ap
40	33	53.2	337	10	US-09-815-242-12539	Sequence 12539, A
41	33	53.2	337	10	US-09-815-242-12876	Sequence 12876, A
42	33	53.2	415	10	US-09-898-461-7	Sequence 7, Appl1
43	33	53.2	419	9	US-10-051-643-125	Sequence 125, App
44	33	53.2	419	9	US-09-880-505-125	Sequence 125, App
45	33	53.2	480	10	US-09-861-270-2	Sequence 2, Appl1

# ALIGNMENTS

RESULT 1  
US-09-764-864-1127  
; Sequence 1127, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: FT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; PRIOR APPLICATION data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1127  
; LENGTH: 146  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (124)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
US-09-764-864-1127

Query Match 62.9%; Score 39; DB 10; Length 146;  
Best Local Similarity 66.7%; Pred. No. 8.1;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ERAEQRLKSOD 12  
|||||  
Db 41 ERAEQRLKSOD 52

RESULT 2  
US-10-101-464A-787  
; Sequence 787, Application US/10101464A  
; Publication No. US20030046728A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Hignins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling

```
FILE REFERENCE: 11000.1020c2
CURRENT APPLICATION NUMBER: US/10/101.464A
CURRENT FILING DATE: 2002-03-18
PRIORITY FILING DATE: 09/704,302
PRIORITY FILING DATE: 2000-11-01
PRIORITY FILING DATE: 09/228,986
PRIORITY FILING DATE: 1999-01-12
PRIORITY FILING DATE: 60/162,866
PRIORITY FILING DATE: 1999-11-01
PRIORITY FILING DATE: PCT/US00/00724
PRIORITY FILING DATE: 2000-01-11
NUMBER OF SEQ ID NOS: 989
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 787
LENGTH: 235
TYPE: PRT
ORGANISM: Pinus radiata
US-10-101-464A-787

Query Match      61.3% Score 38; DB 9; Length 235;
Best Local Similarity 80.0% Pred. No. 20;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy      3 AEOQRLKSOD 12
      1111111111
Db      196 AEOQRLKSOD 205

RESULT 3
US-09-925-299-988
Sequence 988, Application US/09925299
Publication No. US20030040617A9
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIORITY FILING DATE: PCT/US00/05883
PRIORITY FILING DATE: 2000-03-08
PRIORITY FILING DATE: 60/124,270
PRIORITY FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 988
LENGTH: 909
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (32)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (41)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (47)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (48)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (52)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (58)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (62)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (125)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

```
NAME/KEY: SITE
LOCATION: (632)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (851)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-988

Query Match      61.3% Score 38; DB 9; Length 909;
Best Local Similarity 53.8% Pred. No. 88;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Cy      1 ERAEQRLKSODL 13
      1111111111
Db      660 ERAEQRLKSODL 672

RESULT 4
US-09-925-299-988
Sequence 988, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIORITY FILING DATE: PCT/US00/05883
PRIORITY FILING DATE: 2000-03-08
PRIORITY FILING DATE: 60/124,270
PRIORITY FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 988
LENGTH: 909
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (32)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (41)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (47)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (48)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (52)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (58)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (62)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (125)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (851)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-988

Query Match      61.3% Score 38; DB 10; Length 909;
Best Local Similarity 53.8% Pred. No. 88;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

OY 1 ERAEQORLKSODL 13  
| | | | | : | | | | |  
Db 660 EAVEQOKVKMNDL 672

## RESULT 5

US-09-902-941-1932  
; Sequence 1932, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Marnerakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1932  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-902-941-1932

Query Match 59.7%; Score 37; DB 9; Length 404;  
Best Local Similarity 46.2%; Pred. No. 54;  
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 1 ERAEQORLKSODL 13  
| | | | | : | | | | |  
Db 350 DAAERKLSQEI 362

## RESULT 6

US-09-902-941-1934  
; Sequence 1934, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Marnerakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1934  
; LENGTH: 404  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-902-941-1934

Query Match 59.7%; Score 37; DB 9; Length 464;  
Best Local Similarity 46.2%; Pred. No. 63;  
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 1 ERAEQORLKSODL 13  
| | | | | : | | | | |  
Db 410 DAAERKLSQEI 422

## RESULT 7

US-09-854-133-93  
; Sequence 93, Application US/09854133  
; Publication No. US20020183499A1  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.475C10  
; CURRENT APPLICATION NUMBER: US/09/854,133  
; CURRENT FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 735  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 93  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-854-133-93

Query Match 56.5%; Score 35; DB 9; Length 162;  
Best Local Similarity 70.0%; Pred. No. 44;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ERAEQORLKS 10  
| | | | | : | | | | |  
Db 68 EBAEQORLKT 77

## RESULT 8

US-09-738-973-93  
; Sequence 93, Application US/09738973  
; Patent No. US20020110563A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Fling, Steven P.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Elliott, Mark  
; APPLICANT: Mannion, Jane  
; APPLICANT: Kalos, Michael D.  
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.475C9  
; CURRENT APPLICATION NUMBER: US/09/738,973  
; CURRENT FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 587  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 93  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-738-973-93

Query Match 56.5%; Score 35; DB 10; Length 162;  
Best Local Similarity 70.0%; Pred. No. 44;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ERAEQORLKS 10  
| | | | |  
Db 68 EEAEQORLKT 77

RESULT 9  
US-09-986-480-344  
; Sequence 344, Application US/09986480  
; Publication No. US20030027999A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 143 Human Secreted Proteins  
; FILE REFERENCE: PS500P1  
; CURRENT APPLICATION NUMBER: US/09/986,480  
; PRIOR FILING DATE: 2001-11-08  
; PRIOR APPLICATION NUMBER: PCT/US00/12788  
; PRIOR FILING DATE: 2000-05-11  
; PRIOR APPLICATION NUMBER: US 60/134,068  
; PRIOR FILING DATE: 1999-05-13  
; NUMBER OF SEQ ID NOS: 456  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 344  
; LENGTH: 229  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-986-480-344

Query Match 56.5%; Score 35; DB 9; Length 229;  
Best Local Similarity 58.3%; Pred. No. 65;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ERAEQORLKSOD 12  
| | | | |  
Db 139 ERAEQORLKSOD 150

RESULT 10  
US-09-864-761-33475  
; Sequence 33475, Application US/09864761  
; Patent No. US20020046763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aemlica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: CB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 33475  
; LENGTH: 268  
; TYPE: PRT  
; ORGANISM: Homo sapiens

FEATURE:  
; OTHER INFORMATION: MAP TO AF118808.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.87  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.74  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.1  
; OTHER INFORMATION: EST HUMAN HIT: BE295973.1, EVALUE 5.00e-83  
; OTHER INFORMATION: SWISSPROT HIT: P49910, EVALUE 7.00e-38  
US-09-864-761-33475

Query Match 56.5%; Score 35; DB 10; Length 268;  
Best Local Similarity 58.3%; Pred. No. 77;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ERAEQORLKSOD 12  
| | | | |  
Db 13 EELEQORLKSOD 24

RESULT 11  
US-09-916-790-18  
; Sequence 18, Application US/09916790  
; Patent No. US20020061573A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachael  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Silos-Santiago, Immaculada  
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN  
; FILE REFERENCE: KINASE FAMILY MEMBERS AND USBS THEREFOR  
; CURRENT APPLICATION NUMBER: US/09/916,790  
; PRIOR FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: 60/221,543  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Consensus amino acid sequence  
US-09-916-790-18

Query Match 56.5%; Score 35; DB 10; Length 341;  
Best Local Similarity 72.7%; Pred. No. 1e+02;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 AEOORLKSODL 13  
1111 1111  
Db 137 AEOORLKSODL 147

RESULT 12  
US-09-801-574-42  
; Sequence 42, Application US/09801574  
; Patent No. US20020081592A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Peijiang Jeremy  
; APPLICANT: Page, David C.  
; TITLE OF INVENTION: Reproduction-Specific Genes  
; FILE REFERENCE: 0399.2007-002  
; CURRENT APPLICATION NUMBER: US/09/801,574  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: 60/187,518  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: 60/261,557  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-801-574-42

Query Match 56.5%; Score 35; DB 10; Length 458;  
Best Local Similarity 61.5%; Pred. No. 1.4e+02;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ERAOORLKSODL 13  
1111 1111  
Db 11 ELAENSILKNQDL 23

RESULT 13  
US-09-815-242-11532  
; Sequence 11532, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 1410  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11532  
; LENGTH: 529  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-815-242-11532

Query Match 54.8%; Score 34; DB 10; Length 529;  
Best Local Similarity 60.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 AEOORLKSOD 12  
1111 1111  
Db 59 AEEEMKMSQE 68

RESULT 14  
US-09-919-497-59  
; Sequence 59, Application US/09919497  
; Patent No. US20020106662A1  
; GENERAL INFORMATION:  
; APPLICANT: Multer, George L.  
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER  
; FILE REFERENCE: B0801/7225  
; CURRENT APPLICATION NUMBER: US/09/919,497  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/221,735  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 59  
; LENGTH: 767  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-919-497-59

Query Match 54.8%; Score 34; DB 10; Length 767;  
Best Local Similarity 58.3%; Pred. No. 3.6e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ERAOORLKSOD 12  
1111 1111  
Db 538 ERREWSRLKAKD 549

RESULT 15  
US-09-736-457-1819  
; Sequence 1819, Application US/09736457  
; Patent No. US20020168637A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aljun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; PRIOR FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1819  
; LENGTH: 831  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-736-457-1819

Query Match 54.8%; Score 34; DB 9; Length 831;

Best Local Similarity 58.3%; Pred. NO. 3.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 RAQDORLKSODL 13  
||| |:: |||  
Db 25 RAQARVQEDL 36

Search completed: March 13, 2003, 18:54:49  
Job time : 12.1091 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:49:22 ; Search time 8.54545 Seconds  
(without alignments)

53.938 Million cell updates/sec

Title: US-10-050-200-6

Perfect score: 45

Sequence: 1 EKARYLAEM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	88.9	10.9	US-10-057-505-4	Sequence 4, Appl1
2	40	88.9	163	US-09-735-487-8	Sequence 8, Appl1
3	40	88.9	275	US-09-756-551A-4	Sequence 4, Appl1
4	40	88.9	294	US-09-827-822-3	Sequence 3, Appl1
5	40	88.9	492	US-10-003-035-18	Sequence 18, Appl1
6	40	88.9	492	US-09-991-258-5	Sequence 5, Appl1
7	40	88.9	493	US-09-952-060-33	Sequence 33, Appl1
8	40	88.9	500	US-09-968-355-26	Sequence 26, Appl1
9	40	88.9	515	US-09-968-355-20	Sequence 20, Appl1
10	40	88.9	531	US-09-968-355-23	Sequence 23, Appl1
11	40	88.9	583	US-09-968-355-17	Sequence 17, Appl1
12	40	88.9	599	US-10-003-035-59	Sequence 59, Appl1
13	40	88.9	1350	US-09-952-060-35	Sequence 35, Appl1
14	36	80.0	498	US-09-886-156-59	Sequence 59, Appl1
15	36	80.0	498	US-09-886-156-60	Sequence 60, Appl1
16	36	80.0	498	US-09-886-150-59	Sequence 59, Appl1
17	36	80.0	498	US-09-886-150-60	Sequence 60, Appl1
18	36	80.0	498	US-09-886-149-59	Sequence 59, Appl1
19	36	80.0	498	US-09-886-149-60	Sequence 60, Appl1

20	36	80.0	498	9	US-09-886-159-59	Sequence 59, Appl1
21	36	80.0	498	9	US-09-886-159-60	Sequence 60, Appl1
22	34	75.6	219	10	US-09-947-387-44	Sequence 44, Appl1
23	34	75.6	219	10	US-09-893-737-130	Sequence 130, App
24	34	75.6	510	10	US-09-827-822-5	Sequence 5, Appl1
25	34	75.6	521	10	US-09-827-822-4	Sequence 4, Appl1
26	33	73.3	491	10	US-09-997-664-3	Sequence 3, Appl1
27	33	73.3	928	9	US-09-860-846-41	Sequence 41, Appl1
28	33	73.3	928	10	US-09-861-289-41	Sequence 41, Appl1
29	31	68.9	28	9	US-10-127-516-12	Sequence 12, Appl1
30	31	68.9	28	9	US-10-027-629-12	Sequence 12, Appl1
31	31	68.9	28	10	US-09-848-035-10	Sequence 10, Appl1
32	31	68.9	28	10	US-09-815-626-11	Sequence 11, Appl1
33	31	68.9	28	10	US-09-986-224-10	Sequence 10, Appl1
34	31	68.9	511	9	US-09-738-626-5606	Sequence 5606, Ap
35	31	68.9	810	10	US-09-815-242-5586	Sequence 5586, Ap
36	31	68.9	818	10	US-09-815-242-12424	Sequence 12424, A
37	31	68.9	818	10	US-09-815-242-12789	Sequence 12789, A
38	31	68.9	831	10	US-09-815-242-10628	Sequence 10628, A
39	31	68.9	1737	9	US-09-808-602-83	Sequence 83, Appl1
40	31	68.9	2724	9	US-09-808-602-13	Sequence 13, Appl1
41	31	68.9	2733	9	US-09-808-602-8	Sequence 8, Appl1
42	31	68.9	2764	9	US-09-808-602-80	Sequence 80, Appl1
43	31	68.9	2765	9	US-09-808-602-84	Sequence 84, Appl1
44	31	68.9	2802	9	US-09-808-602-81	Sequence 81, Appl1
45	30	66.7	7	9	US-10-112-894-2	Sequence 2, Appl1

#### ALIGNMENTS

RESULT 1  
US-10-057-505-4  
; Sequence 4, Application US/10057505  
; Patent No. US20020164674A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: AURORA BIOSCIENCES CORPORATION  
; APPLICANT: TSIEN, Roger  
; APPLICANT: HEIM, Roger  
; APPLICANT: CUBITT, Andrew  
; TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS  
; FILE REFERENCE: REGEN1260-3  
; CURRENT APPLICATION NUMBER: US/10/057,505  
; CURRENT FILING DATE: 2002-01-25  
; PRIOR APPLICATION NUMBER: US 08/792,553  
; PRIOR FILING DATE: 1997-01-31  
; PRIOR APPLICATION NUMBER: US 09/396,003  
; PRIOR FILING DATE: 1999-09-13  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Linker moiety  
US-10-057-505-4

Query Match 88.9%; Score 40; DB 9; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.011; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 2 KARYLAEM 10  
|||||||  
Db 1 KARYLAEM 9

RESULT 2  
US-09-735-487-8  
; Sequence 8, Application US/09735487  
; Patent No. US20020042679A1  
; GENERAL INFORMATION:

APPLICANT: de BETHUNE, Marie-Pierre  
APPLICANT: HEROGGS, Kurt  
APPLICANT: PAWELIS, Rudi  
TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE  
TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY  
TITLE OF INVENTION: OF HUMAN HIV STRAINS  
FILE REFERENCE: 1377-125P  
CURRENT APPLICATION NUMBER: US/09/735,487  
CURRENT FILING DATE: 2000-12-14  
PRIOR APPLICATION NUMBER: 09/117,217  
PRIOR FILING DATE: 1998-07-24  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 163  
TYPE: PRT  
ORGANISM: HIV-HXB2  
US-09-735-487-8

Query Match 88.9%; Score 40; DB 10; Length 163;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
DB 22 KARVLAEM 30

RESULT 3  
US-09-756-551A-4  
Sequence 4, Application US/09756551A  
Patent No. US20020051768A1  
GENERAL INFORMATION:  
APPLICANT: C. Morrow et al.  
TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL  
TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
TITLE OF INVENTION: USING SAME  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 STATE STREET  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/756,551A  
FILING DATE: 08-JAN-2001  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/376,184  
FILING DATE: 17-AUG-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/987,867  
FILING DATE: 09-DEC-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/389,459  
FILING DATE: 15-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/087,009  
FILING DATE: 01-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauro, Peter C.  
REGISTRATION NUMBER: 32,360  
REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 275 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-756-551A-4

Query Match 88.9%; Score 40; DB 10; Length 275;  
Best Local Similarity 100.0%; Pred. No. 0.53;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
DB 230 KARVLAEM 238

RESULT 4  
US-09-827-822-3  
Sequence 3, Application US/09827822  
Patent No. US20020091086A1  
GENERAL INFORMATION:  
APPLICANT: Anders Vahne  
TITLE OF INVENTION: PEPTIDES THAT BLOCK VIRAL INFECTIVITY  
TITLE OF INVENTION: AND METHODS OF USE THEREOF  
FILE REFERENCE: TRIPEP 003A  
CURRENT APPLICATION NUMBER: US/09/827,822  
CURRENT FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: US/09/370,368  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Human Immunodeficiency Virus Type 1  
US-09-827-822-3

Query Match 88.9%; Score 40; DB 10; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.57;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEM 10  
DB 236 KARVLAEM 244

RESULT 5  
US-10-003-035-18  
Sequence 18, Application US/10003035  
Patent No. US20020155127A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Danher  
TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS  
FILE REFERENCE: 22488-712  
CURRENT APPLICATION NUMBER: US/10/003,035  
CURRENT FILING DATE: 2001-11-01  
PRIOR APPLICATION NUMBER: 09/585,599  
PRIOR FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: PCT/US01/18238  
PRIOR FILING DATE: 2001-06-04  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 18  
LENGTH: 492  
TYPE: PRT  
ORGANISM: Human Immunodeficiency Virus type 1  
US-10-003-035-18

Query Match 88.9%; Score 40; DB 9; Length 492;  
Best Local Similarity 100.0%; Pred. No. 1;



Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
|||||

Db 339 KARVLAEM 347

## RESULT 6

US-09-991-258-5  
; Sequence 5, Application US/09991258  
; Patent No. US20020141975A1  
; GENERAL INFORMATION:  
; APPLICANT: Olmsted, Robert  
; APPLICANT: Keith, Paula  
; APPLICANT: Dryga, Sergey  
; APPLICANT: Caley, Ian  
; APPLICANT: Maughan, Maureen  
; APPLICANT: Johnston, Robert  
; APPLICANT: Davis, Nancy  
; APPLICANT: Swanson, Ronald  
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE  
; FILE REFERENCE: 0113.000103  
; CURRENT APPLICATION NUMBER: US/09/991,258  
; CURRENT FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 09/902,537  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/216,995  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 492  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: NO. US20020141975A1e -  
; OTHER INFORMATION: synthetic construct  
US-09-991-258-5

Query Match 88.9%; Score 40; DB 10; Length 492;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
|||||

Db 356 KARVLAEM 364

## RESULT 7

US-09-952-060-33  
; Sequence 33, Application US/09952060  
; Publication No. US20030044421A1  
; GENERAL INFORMATION:  
; APPLICANT: Emili, Emilio A.  
; APPLICANT: Youli, Rima  
; APPLICANT: Bell, Andrew J.  
; APPLICANT: Chen, Ling  
; APPLICANT: Kaslow, David C.  
; APPLICANT: Shiver, John W.  
; APPLICANT: Toner, Timothy J.  
; APPLICANT: Casimiro, Danilo R.  
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS  
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEE AND  
; FILE REFERENCE: 20747Y  
; CURRENT APPLICATION NUMBER: US/09/952,060  
; CURRENT FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: PCT/US01/28861  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: 60/317,814  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/279,056

; PRIOR FILING DATE: 2001-03-27

; PRIOR APPLICATION NUMBER: 60/233,180

; PRIOR FILING DATE: 2000-09-15

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 33

; LENGTH: 493

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: tPA-gag fusion open reading frame

US-09-952-060-33

Query Match 88.9%; Score 40; DB 9; Length 493;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
|||||

Db 352 KARVLAEM 360

## RESULT 8

US-09-968-355-26  
; Sequence 26, Application US/09968355  
; Patent No. US20020094523A1  
; GENERAL INFORMATION:  
; APPLICANT: Sakalian, Michael  
; APPLICANT: Hunter, Eric  
; TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays  
; FILE REFERENCE: UAB-100XCI  
; CURRENT APPLICATION NUMBER: US/09/968,355  
; CURRENT FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: 60/236,273  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1)-(500)  
; OTHER INFORMATION: HIV (FS-) GAG amino acid sequence  
US-09-968-355-26

Query Match 88.9%; Score 40; DB 10; Length 500;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10  
|||||

Db 359 KARVLAEM 367

## RESULT 9

US-09-968-355-20  
; Sequence 20, Application US/09968355  
; Patent No. US20020094523A1  
; GENERAL INFORMATION:  
; APPLICANT: Sakalian, Michael  
; APPLICANT: Hunter, Eric  
; TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays  
; FILE REFERENCE: UAB-100XCI  
; CURRENT APPLICATION NUMBER: US/09/968,355  
; CURRENT FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: 60/236,273  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 515

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(515)
; OTHER INFORMATION: Ch3a amino acid sequence
US-09-968-355-20
```

```

Query Match      88.9%; Score 40; DB 10; Length 515;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 KARVLAEM 10
        |||||
Db       359 KARVLAEM 367
```

```

RESULT 10
US-09-968-355-23
; Sequence 23, Application US/09968355
; Patent No. US20020094523A1
; GENERAL INFORMATION:
; APPLICANT: Sakalian, Michael
; APPLICANT: Hunter, Eric
; TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays
; FILE REFERENCE: UAB-100XC1
; CURRENT APPLICATION NUMBER: US/09/968,355
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,273
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 23
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(531)
; OTHER INFORMATION: Ch3 amino acid sequence
US-09-968-355-23
```

```

Query Match      88.9%; Score 40; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 KARVLAEM 10
        |||||
Db       359 KARVLAEM 367
```

```

RESULT 11
US-09-968-355-17
; Sequence 17, Application US/09968355
; Patent No. US20020094523A1
; GENERAL INFORMATION:
; APPLICANT: Sakalian, Michael
; APPLICANT: Hunter, Eric
; TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays
; FILE REFERENCE: UAB-100XC1
; CURRENT APPLICATION NUMBER: US/09/968,355
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,273
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 17
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(583)
```

```

; OTHER INFORMATION: Chimera 4 amino acid sequence
US-09-968-355-17
```

```

Query Match      88.9%; Score 40; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 KARVLAEM 10
        |||||
Db       359 KARVLAEM 367
```

```

RESULT 12
US-10-003-035-59
; Sequence 59, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 59
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Gag-PI
US-10-003-035-59
```

```

Query Match      88.9%; Score 40; DB 9; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 KARVLAEM 10
        |||||
Db       359 KARVLAEM 367
```

```

RESULT 13
US-09-952-060-35
; Sequence 35, Application US/09952060
; Publication No. US20030044421A1
; GENERAL INFORMATION:
; APPLICANT: Emini, Emilio A.
; APPLICANT: Youll, Rima
; APPLICANT: Betz, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Shiver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317,814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279,056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 1350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Codon optimized gag-1A pol fusion
US-09-952-060-35

Query Match
Best Local Similarity 88.9%; Score 40; DB 9; Length 1350;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10
DB 359 KARVLAEM 367

RESULT 14
US-09-886-156-59
; Sequence 59, Application US/09886156
; Patent No. US20020155428A1
; GENERAL INFORMATION:
; APPLICANT: Guerlier, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,156
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
US-09-886-156-59

Query Match
Best Local Similarity 80.0%; Score 36; DB 9; Length 498;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10
DB 358 KARVLAEM 366

RESULT 15
US-09-886-156-60
; Sequence 60, Application US/09886156
; Patent No. US20020155428A1
; GENERAL INFORMATION:
; APPLICANT: Guerlier, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,156
; CURRENT FILING DATE: 2001-06-22
```

```
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
US-09-886-156-60
```

```
Query Match
Best Local Similarity 80.0%; Score 36; DB 9; Length 498;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEM 10
DB 358 KARVLAEM 366
```

Search completed: March 13, 2003, 18:54:48  
Job time : 9.54545 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:49:22 ; Search time 8.54545 seconds  
(without alignments)  
53.938 Million cell updates/sec

Title: US-10-050-200-5  
Perfect score: 44  
Sequence: 1 EKARVLAEEA 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA:\*  
1: /cgn2\_6/prodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/prodata/1/pubppa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/1/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/1/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/1/pubppa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/1/pubppa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/prodata/1/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/1/pubppa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/prodata/1/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/1/pubppa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/prodata/1/pubppa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/prodata/1/pubppa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	86.4	219	10 US-09-893-737-130	Sequence 130, App
2	35	79.5	109	10 US-10-057-505-4	Sequence 4, Appl1
3	35	79.5	163	10 US-09-735-487-8	Sequence 4, Appl1
4	35	79.5	275	10 US-09-756-551A-4	Sequence 4, Appl1
5	35	79.5	294	10 US-09-827-822A-3	Sequence 4, Appl1
6	35	79.5	492	9 US-10-003-035-18	Sequence 18, Appl1
7	35	79.5	492	10 US-09-991-258-5	Sequence 5, Appl1
8	35	79.5	493	9 US-08-952-060-33	Sequence 33, Appl1
9	35	79.5	500	10 US-09-968-355-26	Sequence 26, Appl1
10	35	79.5	515	10 US-09-968-355-20	Sequence 20, Appl1
11	35	79.5	531	10 US-09-968-355-23	Sequence 23, Appl1
12	35	79.5	583	10 US-09-968-355-17	Sequence 17, Appl1
13	35	79.5	599	9 US-10-003-035-59	Sequence 59, Appl1
14	35	79.5	1350	9 US-08-952-060-35	Sequence 35, Appl1
15	35	79.5	928	9 US-08-860-846-41	Sequence 41, Appl1
16	33	75.0	928	10 US-09-861-289-41	Sequence 41, Appl1
17	32	72.7	135	10 US-09-864-761-48619	Sequence 48619, A
18	32	72.7	454	10 US-09-815-242-5209	Sequence 5209, Ap
19	32	72.7	491	10 US-09-997-664-3	Sequence 3, Appl1

20	32	72.7	510	10 US-09-827-822-5	Sequence 5, Appl1
21	32	72.7	521	10 US-09-827-822-4	Sequence 4, Appl1
22	31	70.5	498	9 US-09-886-156-59	Sequence 59, Appl1
23	31	70.5	498	9 US-09-886-156-60	Sequence 60, Appl1
24	31	70.5	498	9 US-09-886-150-59	Sequence 59, Appl1
25	31	70.5	498	9 US-09-886-150-60	Sequence 60, Appl1
26	31	70.5	498	9 US-09-886-149-59	Sequence 59, Appl1
27	31	70.5	498	9 US-09-886-149-60	Sequence 60, Appl1
28	31	70.5	498	9 US-09-886-159-59	Sequence 59, Appl1
29	31	70.5	498	9 US-09-886-159-60	Sequence 60, Appl1
30	31	70.5	1148	10 US-09-815-242-13826	Sequence 13826, A
31	31	70.5	1737	9 US-09-808-602-83	Sequence 83, Appl1
32	31	70.5	2724	9 US-09-808-602-13	Sequence 13, Appl1
33	31	70.5	2733	9 US-09-808-602-8	Sequence 8, Appl1
34	31	70.5	2764	9 US-09-808-602-80	Sequence 80, Appl1
35	31	70.5	2765	9 US-09-808-602-84	Sequence 84, Appl1
36	31	70.5	2802	9 US-09-808-602-81	Sequence 81, Appl1
37	30	68.2	7	9 US-10-112-894-2	Sequence 2, Appl1
38	30	68.2	7	10 US-09-947-387-29	Sequence 29, Appl1
39	30	68.2	285	10 US-09-764-898-187	Sequence 187, App
40	30	68.2	251	9 US-09-738-626-5378	Sequence 5378, App
41	30	68.2	326	10 US-09-841-132-447	Sequence 447, App
42	30	68.2	445	10 US-09-815-242-10316	Sequence 10316, A
43	30	68.2	445	10 US-09-815-242-13762	Sequence 13762, A
44	30	68.2	532	10 US-09-801-368-360	Sequence 360, App
45	30	68.2	538	10 US-09-827-822-9	Sequence 9, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-893-737-130  
Sequence 130, Application US/09893737  
Patent No. US20020110855A1  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS  
FILE REFERENCE: 00-41  
CURRENT APPLICATION NUMBER: US/09/893,737  
CURRENT FILING DATE: 2001-06-28  
PRIOR APPLICATION NUMBER: US 60/215,446  
PRIOR FILING DATE: 2000-06-30  
NUMBER OF SEQ ID NOS: 329  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 130  
LENGTH: 219  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-893-737-130

Query Match 86.4%; Score 38; DB 10; Length 219;  
Best Local Similarity 80.0%; Pred. No. 1.5;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKARVLAEEA 10  
||:|||||  
Db 52 EKSRILAEEA 61

RESULT 2  
US-10-057-505-4  
Sequence 4, Application US/10057505  
Patent No. US20020164674A1  
GENERAL INFORMATION:  
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: AUTORA BIOSCIENCES CORPORATION  
APPLICANT: TSIEN, Roger  
APPLICANT: HEIM, Roger  
APPLICANT: CUBITT, Andrew  
TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS  
FILE REFERENCE: REGEN1260-3

```

; CURRENT APPLICATION NUMBER: US/10/057,505
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 08/792,553
; PRIOR FILING DATE: 1997-01-31
; PRIOR APPLICATION NUMBER: US 09/396,003
; PRIOR FILING DATE: 1999-09-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Linker moiety
US-10-057-505-4

```

```

Query Match          79.5%; Score 35; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 KARVILAEA 9
        |||||||
Db      1 KARVILAEA 8

```

```

RESULT 3
US-09-735-487-8
; Sequence 8, Application US/09735487
; Patent No. US20020042679A1
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUMELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/735,487
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,217
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 163
; TYPE: PRT
; ORGANISM: HIV-HXB2
US-09-735-487-8

```

```

Query Match          79.5%; Score 35; DB 10; Length 163;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 KARVILAEA 9
        |||||||
Db      22 KARVILAEA 29

```

```

RESULT 4
US-09-736-551A-4
; Sequence 4, Application US/09756551A
; Patent No. US20020051768A1
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS

```

```

; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/756,551A
; FILING DATE: 08-JAN-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/376,184
; FILING DATE: 17-AUG-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/987,867
; FILING DATE: 09-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/389,459
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauro, Peter C.
; REGISTRATION NUMBER: 32,360
; REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-756-551A-4

```

```

Query Match          79.5%; Score 35; DB 10; Length 275;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 KARVILAEA 9
        |||||||
Db      230 KARVILAEA 237

```

```

RESULT 5
US-09-827-822-3
; Sequence 3, Application US/09827822
; Patent No. US20020091086A1
; GENERAL INFORMATION:
; APPLICANT: Anders Vahine
; TITLE OF INVENTION: PEPTIDES THAT BLOCK VIRAL INFECTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP 003A
; CURRENT APPLICATION NUMBER: US/09/827,822
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US/09/370,368
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO 3
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus Type 1
US-09-827-822-3

```

```

Query Match          79.5%; Score 35; DB 10; Length 294;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 2 KARVLAEA 9  
|||||||  
DB 236 KARVLAEA 243

## RESULT 6

US-10-003-035-18  
; Sequence 18, Application US/10003035  
; Patent No. US20020135127A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Danher  
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS  
; FILE REFERENCE: 22488-712  
; CURRENT APPLICATION NUMBER: US/10/003,035  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 09/585,599  
; PRIOR FILING DATE: 2000-06-02  
; PRIOR APPLICATION NUMBER: PCT/US01/18238  
; PRIOR FILING DATE: 2001-06-04  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 492  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-003-035-18

Query Match 79.5%; Score 35; DB 9; Length 492;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
|||||||  
DB 339 KARVLAEA 346

## RESULT 7

US-09-991-258-5  
; Sequence 5, Application US/09991258  
; Patent No. US20020141975A1  
; GENERAL INFORMATION:  
; APPLICANT: Olmsted, Robert  
; APPLICANT: Keith, Paula  
; APPLICANT: Dryga, Sergey  
; APPLICANT: Caley, Ian  
; APPLICANT: Maughan, Maureen  
; APPLICANT: Johnston, Robert  
; APPLICANT: Davis, Nancy  
; APPLICANT: Swanstrom, Ronald  
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE  
; FILE REFERENCE: 0113.000103  
; CURRENT APPLICATION NUMBER: US/09/991,258  
; CURRENT FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 09/902,537  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/216,995  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 492  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =  
; OTHER INFORMATION: synthetic construct  
US-09-991-258-5

Query Match 79.5%; Score 35; DB 10; Length 492;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
|||||||  
DB 356 KARVLAEA 363

## RESULT 8

US-09-952-060-33  
; Sequence 33, Application US/09952060  
; Publication No. US20030044421A1  
; GENERAL INFORMATION:  
; APPLICANT: Emili, Emilio A.  
; APPLICANT: Youli, Rima  
; APPLICANT: Belt, Andrew J.  
; APPLICANT: Chen, Ling  
; APPLICANT: Kaslow, David C.  
; APPLICANT: Shiver, John W.  
; APPLICANT: Toner, Timothy J.  
; APPLICANT: Casimiro, Danilo R.  
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS  
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND  
; FILE REFERENCE: 20747Y  
; CURRENT APPLICATION NUMBER: US/09/952,060  
; CURRENT FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: PCT/US01/28861  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: 60/317,814  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/279,056  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/233,180  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 33  
; LENGTH: 493  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: tPA-gag fusion open reading frame  
US-09-952-060-33

Query Match 79.5%; Score 35; DB 9; Length 493;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KARVLAEA 9  
|||||||  
DB 352 KARVLAEA 359

## RESULT 9

US-09-968-355-26  
; Sequence 26, Application US/09968355  
; Patent No. US20020094523A1  
; GENERAL INFORMATION:  
; APPLICANT: Sakalian, Michael  
; APPLICANT: Hunter, Eric  
; TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays  
; FILE REFERENCE: UAB-100XCI  
; CURRENT APPLICATION NUMBER: US/09/968,355  
; CURRENT FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: 60/236,273  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE

Query Match 79.5%; Score 35; DB 10; Length 493;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCATION: (1)..(500)  
OTHER INFORMATION: HIV (FS-) Gag amino acid sequence  
US-09-968-355-26

Query Match 79.5%; Score 35; DB 10; Length 500;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEA 9  
|||||  
DB 359 KARVLAEA 366

RESULT 10  
US-09-968-355-20  
Sequence 20, Application US/09968355  
Patent No. US20020094523A1  
GENERAL INFORMATION:  
APPLICANT: Sakalian, Michael  
APPLICANT: Hunter, Eric  
TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays  
FILE REFERENCE: UAB-100XC1  
CURRENT APPLICATION NUMBER: US/09/968,355  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: 60/236,273  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 20  
LENGTH: 515  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (1)..(515)  
OTHER INFORMATION: Ch3a amino acid sequence  
US-09-968-355-20

Query Match 79.5%; Score 35; DB 10; Length 515;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEA 9  
|||||  
DB 359 KARVLAEA 366

RESULT 11  
US-09-968-355-23  
Sequence 23, Application US/09968355  
Patent No. US20020094523A1  
GENERAL INFORMATION:  
APPLICANT: Sakalian, Michael  
APPLICANT: Hunter, Eric  
TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays  
FILE REFERENCE: UAB-100XC1  
CURRENT APPLICATION NUMBER: US/09/968,355  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: 60/236,273  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 23  
LENGTH: 531  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (1)..(531)  
OTHER INFORMATION: Ch3 amino acid sequence  
US-09-968-355-23

Query Match 79.5%; Score 35; DB 10; Length 531;

Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEA 9  
|||||  
DB 359 KARVLAEA 366

RESULT 12  
US-09-968-355-17  
Sequence 17, Application US/09968355  
Patent No. US20020094523A1  
GENERAL INFORMATION:  
APPLICANT: Sakalian, Michael  
APPLICANT: Hunter, Eric  
TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays  
FILE REFERENCE: UAB-100XC1  
CURRENT APPLICATION NUMBER: US/09/968,355  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: 60/236,273  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 583  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (1)..(583)  
OTHER INFORMATION: Ch4mera 4 amino acid sequence  
US-09-968-355-17

Query Match 79.5%; Score 35; DB 10; Length 583;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEA 9  
|||||  
DB 359 KARVLAEA 366

RESULT 13  
US-10-003-035-59  
Sequence 59, Application US/10003035  
Patent No. US20020155127A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Danher  
TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS  
FILE REFERENCE: 22488-712  
CURRENT APPLICATION NUMBER: US/10/003,035  
PRIOR FILING DATE: 2001-11-01  
PRIOR APPLICATION NUMBER: 09/585,599  
PRIOR FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: PCT/US01/18238  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 59  
LENGTH: 599  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Gag-P1  
US-10-003-035-59

Query Match 79.5%; Score 35; DB 9; Length 599;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KARVLAEA 9  
|||||  
DB 359 KARVLAEA 366



## RESULT 14

US-09-952-060-35  
; Sequence 35, Application US/09952060  
; Publication No. US20030044421A1  
; GENERAL INFORMATION:  
; APPLICANT: Eminli, Emilio A.  
; APPLICANT: Youll, Rima  
; APPLICANT: Belt, Andrew J.  
; APPLICANT: Chen, Ling  
; APPLICANT: Kaslow, David C.  
; APPLICANT: Shiver, John W.  
; APPLICANT: Toner, Timothy J.  
; APPLICANT: Casimiro, Danilo R.  
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS  
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND  
; FILE REFERENCE: 20747Y  
; CURRENT APPLICATION NUMBER: US/09/952,060  
; CURRENT FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: PCT/US01/28861  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: 60/317,814  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/279,056  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/233,180  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 1350  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Codon optimized gag-1A pol fusion  
US-09-952-060-35

## Query Match

79.5%; Score 35; DB 9; Length 1350;

Best Local Similarity 100.0%; Pred. No. 50;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 2 KARYLAEA 9  
|||||||

Db 359 KARYLAEA 366

## RESULT 15

US-09-860-846-41  
; Sequence 41, Application US/09860846  
; Patent No. US20020164742A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600,438US1  
; CURRENT APPLICATION NUMBER: US/09/860,846  
; CURRENT FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 41  
; LENGTH: 928  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-860-846-41

Query Match 75.0%; Score 33; DB 9; Length 928;  
Best Local Similarity 77.8%; Pred. No. 84;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EKARYLAEA 9  
:|||||||

Db 739 OKARYLAEA 747

Search completed: March 13, 2003, 18:54:47  
Job time : 9.54545 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 18:45:22 ; Search time 22.8 Seconds  
(without alignments)  
99.409 Million cell updates/sec

Title: US-10-050-200-3  
Perfect score: 50  
Sequence: 1 KELAELRESTS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008  
Listing first 405 summaries

Database :

SPREMBL.21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp\_uniclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	82.0	410	4 Q96LM3	Q96LM3 homo sapien
2	41	82.0	1663	4 Q9H0J4	Q9H0J4 homo sapien
3	38	76.0	131	11 Q99PH8	Q99PH8 mus musculi
4	37	74.0	264	16 Q8XUJ1	Q8XUJ1 ralatonia s
5	36	72.0	169	11 Q9WV90	Q9WV90 marimota mon
6	36	72.0	265	10 Q9LCE3	Q9LCE3 streptomyces
7	36	72.0	284	16 Q9RL36	Q9RL36 streptomyces
8	36	72.0	352	3 Q94I35	Q94I35 streptomyces
9	36	72.0	614	16 Q9LCE9	Q9LCE9 streptomyces
10	36	72.0	876	10 Q855V4	Q855V4 oryza sativ
11	36	72.0	1713	5 Q9V3Q9	Q9V3Q9 drosophila
12	35	70.0	120	10 Q940A5	Q940A5 arabidopsis
13	35	70.0	317	5 Q97338	Q97338 physarum po
14	35	70.0	436	16 Q9H050	Q9H050 pseudomonas
15	35	70.0	482	5 Q9VNM2	Q9VNM2 drosophila
16	35	70.0	588	5 Q9W5J8	Q9W5J8 drosophila

17	35	70.0	625	16 Q83666	Q83666 treponema p
18	35	70.0	1140	10 Q9LYV5	Q9LYV5 arabidopsis
19	35	70.0	1486	5 Q967Y2	Q967Y2 drosophila
20	35	70.0	1486	5 Q95RE5	Q95RE5 drosophila
21	35	70.0	1486	5 Q9V9D3	Q9V9D3 drosophila
22	34	68.0	74	17 Q9H0U2	Q9H0U2 halobacteri
23	34	68.0	94	9 Q8MET7	Q8MET7 bacterioph
24	34	68.0	159	17 Q85626	Q85626 pyrococcus
25	34	68.0	281	16 Q8RDQ3	Q8RDQ3 fusobacteri
26	34	68.0	294	5 Q9TZB0	Q9TZB0 caenorhabdi
27	34	68.0	324	10 Q94I10	Q94I10 oryza sativ
28	34	68.0	328	10 Q82595	Q82595 arabidopsis
29	34	68.0	370	4 Q96PW6	Q96PW6 homo sapien
30	34	68.0	377	3 Q8RTW7	Q8RTW7 aspergillus
31	34	68.0	394	10 Q94I12	Q94I12 oryza sativ
32	34	68.0	508	16 Q98EP9	Q98EP9 rhizobium l
33	34	68.0	568	17 Q97WX9	Q97WX9 sulfolobus
34	34	68.0	834	10 Q8W056	Q8W056 oryza sativ
35	34	68.0	1604	13 Q90662	Q90662 gallus gall
36	34	68.0	2148	5 Q9BDJ3	Q9BDJ3 physarum po
37	34	68.0	2228	5 Q60991	Q60991 plasmodium
38	33	66.0	2228	5 Q60991	Q60991 plasmodium
39	33	66.0	236	2 Q47610	Q47610 escherichia
40	33	66.0	84	15 Q902K1	Q902K1 human immu
41	33	66.0	96	11 Q60729	Q60729 mus musculi
42	33	66.0	141	17 Q803A1	Q803A1 pyrococcus
43	33	66.0	188	11 Q04365	Q04365 mus musculi
44	33	66.0	216	17 Q8ZUH2	Q8ZUH2 pyrobaculum
45	33	66.0	267	16 Q9RV73	Q9RV73 deinococcus

## ALIGNMENTS

### RESULT 1

Q96LM3	PRELIMINARY:	PRT:	410 AA.
AC Q96LM3;			
DT 01-DEC-2001 (TREMBLrel. 19, Created)			
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE CDNA FLJ25373 f1s, clone TST01951.			
OS Homo sapiens (human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX NCBI_Taxid=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
KC TISSUE=TESTIS;			
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,			
RA Hotuta T., Hirooka S., Murakawa K., Takiguchi S., Kusano J.,			
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chida Y.,			
RA Suzuki Y., Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,			
RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,			
RA Kawakami B., Nagai K., Isogai T., Sugano S.;			
RT "MEDO human cDNA sequencing project."			
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
DR EMBL: AK058102; BAB71667.1;			
SQ SEQUENCE 410 AA; 46072 MW; D215FD128A0BD093 CRC64;			

Query Match 82.0%; Score 41; DB 4; Length 410;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELAELRES 9  
|||||||

DB 122 KELAELRES 130

### RESULT 2

Q9H0J4 PRELIMINARY: PRT: 1663 AA.  
AC Q9H0J4;

DT 01-MAR-2001 (TRENBLREL. 16, Created)  
 DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)  
 DE Hypothetical protein 180.8 kDa protein.  
 GN DKFZP434P0316.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-TESTIS;  
 RX MEDLINE-21154917; PubMed-11230166;  
 RA Wieman S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,  
 RA Ansgore W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,  
 RA Meves H.W., Osterwaelde B., Obermaier B., Tampe J., Heubner D.,  
 RA Wamutke R., Korn B., Klein M., Poustka A.;  
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and  
 RT Analysis of 500 Novel Complete Protein Coding Human cDNAs."  
 RL Genome Res. 11:422-435(2001).  
 DR EMBL: AL136774; CAB66708.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1663 AA; 180826 MW; 62E141FEF5C4DBE CRC64;

Query Match  
 Best Local Similarity 82.0%; Score 41; DB 4; Length 1663;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAELRES 9  
 DB 1141 KELAELRES 1149

RESULT 3  
 ID Q99PH8 PRELIMINARY; PRT; 131 AA.  
 AC Q99PH8;  
 DT 01-JUN-2001 (TRENBLREL. 17, Created)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)  
 DE Fas ligand.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RA Zeylun A., Nagarkatti M., Nagarkatti P.S.;  
 RT "Isolation and characterization of a new naturally occurring variant of  
 RT Fas ligand that is functional and expressed only in membrane bound  
 RT form."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF286572; AAG60016.1; -.  
 SQ SEQUENCE 131 AA; 14538 MW; 79B9C3E7242EB424 CRC64;

Query Match  
 Best Local Similarity 76.0%; Score 38; DB 11; Length 131;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KELAELRES 10  
 DB 107 KELAELRES 116

RESULT 4  
 ID Q8XU01 PRELIMINARY; PRT; 264 AA.  
 AC Q8XU01;  
 DT 01-MAR-2002 (TRENBLREL. 20, Created)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)

DE Hypothetical protein RSC3094.  
 GN RSC3094 OR RS00547.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GM11000;  
 RX MEDLINE-21681879; PubMed-11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arelat M., Billault A., Brottier P., Camu J.C., Cartolico L.,  
 RA Chandler M., Choisme N., Claudel-Renard C., Cunha S., Demange N.,  
 RA Gaspin C., Lavie W., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Sigler P., Trebault P., Whalen M., Winkler P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT Genome sequence of the plant pathogen *Ralstonia solanacearum*.  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646073; CAD16803.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 264 AA; 28781 MW; F68F861312DD3F61 CRC64;

Query Match  
 Best Local Similarity 74.0%; Score 37; DB 16; Length 264;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 KELAELRES 11  
 DB 77 KELAELRES 85

RESULT 5  
 ID Q9WV90 PRELIMINARY; PRT; 169 AA.  
 AC Q9WV90;  
 DT 01-NOV-1999 (TRENBLREL. 12, Created)  
 DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE Fas ligand (Fragment).  
 OS Marmota monax (Woodchuck).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;  
 OC Marmota.  
 OX NCBI\_TaxID=9995;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEALTHY LIVER;  
 RX MEDLINE-20010026; PubMed-10540161;  
 RA Hodgson P.D., Grant M.D., Michalak T.I.;  
 RT "Perforin and Fas/Fas ligand-mediated cytotoxicity in acute and  
 RT chronic woodchuck viral hepatitis."  
 RL Clin. Exp. Immunol. 118:63-70(1999).  
 DR EMBL: AF152368; AAD38387.1; -.  
 DR HSP; P50591; IDAV.  
 DR Interpro; IPR003636; TNF-abc.  
 DR Interpro; IPR00478; TNF-family.  
 DR Pfam; PF00229; TNF; 1.  
 DR Prodom; P002012; TNF-abc; 1.  
 DR SMART; SM00207; TNF; 1.  
 DR PROSITE; PS00251; TNF\_1; 1.  
 DR PROSITE; PS50049; TNF\_2; 1.  
 FT NON\_TER 169  
 FT NON\_TER 169  
 SQ SEQUENCE 169 AA; 19274 MW; FDE395B014717B6B CRC64;

Query Match  
 Best Local Similarity 72.0%; Score 36; DB 11; Length 169;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KELAELRES 9  
 DB 6 KELAELRES 14

```

RESULT 6
09LJEB3 ID 09LJEB3 PRELIMINARY; PRT; 265 AA.
AC 09LJEB3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE Genomic DNA, chromosome 3, pl clone: MRP15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatopsida; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=9702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Kaneo T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,
RT TAC and BAC clones."
RL DNA Res. 7:217-221(2000).
DR EMBL; AP000603; BAB01755.1;
SO SEQUENCE 265 AA; 30326 MW; DC3A51749666728F CRC64;

Query Match 72.0%; Score 36; DB 10; Length 265;
Best Local Similarity 72.7%; Pred. No. 53;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KELAELRES 11
DB 131 KELAELRES 141

RESULT 7
09RLJ6 ID 09RLJ6 PRELIMINARY; PRT; 284 AA.
AC 09RLJ6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical protein SC00451.
GN SC00451 OR SCF51A.29C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Murphy L., Harris D.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Thompson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA MEDLINE=97000351; PubMed=8843436;
RA Reidenbach M., Krieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
RN [4]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Krieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Krieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL; AL121596; CAB56681.1;
DR InterPro; IPR001932; PP2C-like.
DR SMART; SM00332; PP2C; 1.
DR SMART; SM00331; PP2C-SIG; 1.
KW Hypothetical protein
SO SEQUENCE 284 AA; 29620 MW; C80B71C89070FAD CRC64;

Query Match 72.0%; Score 36; DB 16; Length 284;
Best Local Similarity 88.9%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KELAELRES 9
DB 25 KELAELRES 33

RESULT 8
094135 ID 094135 PRELIMINARY; PRT; 352 AA.
AC 094135;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Ketol-acid reductoisomerase (EC 1.1.1.86).
GN AHR1.
GN PIROMYCES SP. E2.
OS Eukaryota; Fungi; Chytridiomycota; Neocallimastix;
OC Neocallimastixaceae; Piromyces.
OX NCBI_TaxID=73868;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-E2;
RX MEDLINE=99141007; PubMed=9988478;
RA Akhmanova A., Voncken F., Harhangi H., Hosea K.M., Vogels G.D.,
RA Hackstein J.H.P.;
RT "Cytosolic enzymes with a mitochondrial ancestry from the anaerobic
RT chytrid Piromyces sp. E2."
RL Mol. Microbiol. 30:1017-1027(1998).
CC -1- CATALYTIC ACTIVITY: 2,3-DIHYDROXYISOVALERATE + NADP(+) = 2-
CC ACETOLACTATE + NADPH (ALSO: 2,3-DIHYDROXYACID-3-METHYLVALERATE +
CC NADP(+) = 2-ACETO-2-HYDROXYBUTYRATE + NADPH).
DR EMBL; Y16743; CAA76356.1;
DR HSSP; Q01292; IOMG.
DR InterPro; IPR000506; ACH_Isomrctse.
DR Pfam; PF01450; IIVC; 1.
DR TIGRfams; TIGR00465; IIVC; 1.
KW Isomerase; NADP; Oxidoreductase.
SO SEQUENCE 352 AA; 39453 MW; BE2056BC934C1C14 CRC64;

Query Match 72.0%; Score 36; DB 3; Length 352;
Best Local Similarity 88.9%; Pred. No. 72;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KELAELRES 9
DB 322 KELAELRES 330

RESULT 9

```

0912P9  
ID Q912P9 PRELIMINARY: PRT: 614 AA.  
AC Q912P9;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Hypothetical protein SC02524.  
GN SC02524 OR SC121.27C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycetaceae; Streptomycetes.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2) / M145;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Godle A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,  
RA Hopwood D.A.,  
RT Complete genome sequence of the model actinomycete Streptomyces  
RT coelicolor A3(2)."  
RL Nature 417:141-147(2002).  
DR EMBL: AL137166; CAB69745.1;  
KW Hypothetical protein.  
SQ SEQUENCE 614 AA; 68702 MW; DFB8B38EC2368BB CRC64;

Query Match 72.0%; Score 36; DB 16; Length 614;  
Best Local Similarity 63.6%; Pred. No. 1.3e+02;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11  
DB 161 KAIAREXTN 171

RESULT 10  
OBS5V4 PRELIMINARY: PRT: 876 AA.  
AC OBS5V4;  
DT 01-JUN-2002 (TREMblrel. 21, Created)  
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Hypothetical 97.8 kDa protein.  
GN OJ1015P07.13.  
OS Oryza sativa (Japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
RA Sasaki C., Henry D., Oates R., Simmons J.;  
RT "Rice Genomic Sequence."  
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC104427; AAM19115.1;  
KW Hypothetical protein.  
SQ SEQUENCE 876 AA; 97835 MW; 26AE76D87C3E0180 CRC64;

Query Match 72.0%; Score 36; DB 10; Length 876;  
Best Local Similarity 72.7%; Pred. No. 1.9e+02;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KELAIRESTS 11  
DB 253 KELAIRESTS 263

RESULT 11  
O9V309 PRELIMINARY: PRT: 1713 AA.  
AC O9V309;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE CG15274 protein.  
GN GABA-B-R1 OR BG:DS00929.6 OR CG15274.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkovec D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,  
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
RN [12]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE=99403001; PubMed=10471707;  
RA Ashburner M., Mista S., Roote J., Lewis S.E., Blazek R., Davis T.,  
RA Doyle C., Galie R., George R., Harris N., Hartwell G., Harvey D.,  
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,  
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K.,  
RA Ceiniker S., Rubin G.M.;  
RT "An exploration of the sequence of a 2.9-Mb region of the genome of  
RT Drosophila melanogaster: the Adh region."  
RL Genetics 153:179-219(1999).  
RN [13]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RA Ceiniker S.E., Agbayani A., Arcalena T.T., Baxter E., Blazek R.G.,  
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,

RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,  
 RA Houston K.A., Hummasti S.R., Karia K., Kearney L., Kim E., Lee B.,  
 RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi A.,  
 RA Nixon K., Pacled J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,  
 RA Setti H., Shit E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,  
 RA Zieran L.L., Rubin G.M.,  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003646; AAF53431.1; -;  
 DR EMBL: AE003411; AAF44910.1; -;  
 DR FlyBase: FBgn0028824; GABA-B-RL.  
 DR InterPro: IPR000337; GPCR\_Mgr.  
 DR InterPro: IPR000402; Na/K-ATPase\_beta.  
 DR Pfam: PF00003; 7tm\_3; 1.  
 DR Pfam: PF00287; Na/K-ATPase; 1.  
 DR PROSITE: PS50259; G\_PROTEIN\_RECEP\_F3\_4; 2.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1713 AA; 19513 MW; 8474BEC5F9A27AC CRC64;

Query Match 72.0%; Score 36; DB 5; Length 1713;  
 Best Local Similarity 72.7%; Pred. No. 3; 7e+02;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KELAURESTS 11  
 DB 857 KELAURESTN 867

RESULT 12  
 ID 0940A5 PRELIMINARY; PRT; 120 AA.  
 AC 0940A5;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Hypothetical 13.5 kDa protein.  
 GN AT1G06525.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,  
 RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,  
 RA Yu S., Bowser L., Carlinici P., Chen H., Cheuk R., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,  
 RA Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,  
 RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,  
 RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.,  
 RT Full length cDNA of gene At1g06525.  
 RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY056146; AAL07225.1; -;  
 DR InterPro: IPR001230; Prenyl\_site.  
 DR PROSITE: PS00294; PRENYLATION; UNKNOWN\_1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 120 AA; 13546 MW; DD4B45F932C66C33 CRC64;

Query Match 70.0%; Score 35; DB 10; Length 120;  
 Best Local Similarity 87.5%; Pred. No. 37;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELAURE 8  
 DB 91 KELAURE 98

RESULT 13  
 ID 097338 PRELIMINARY; PRT; 317 AA.  
 AC 097338;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE Putative type II myosin heavy chain (Fragment).  
 GN MYND.  
 OS Physarum polycephalum (slime mold).  
 OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;  
 OC Physarum.  
 OX NCBI\_TaxID=5791;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CL;  
 RX MEDLINE=20059328; Pubmed-10591846;  
 RA Bailey J., Cook L.J., Kilmer-Barber R., Swanson E.,  
 RA Solonica-Krezel L., Lohman K.N., Dove W.F., Dee J., Anderson R.W.,  
 RT Identification of three genes expressed primarily during development  
 in Physarum polycephalum.  
 RL Arch. Microbiol. 172:364-376(1999).  
 DR EMBL: AJ133501; CAB39170.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 317 AA; 36825 MW; 486504902049302C CRC64;

Query Match 70.0%; Score 35; DB 5; Length 317;  
 Best Local Similarity 70.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KELAUREST 10  
 DB 144 KELAUREKS 153

RESULT 14  
 ID 09H050 PRELIMINARY; PRT; 436 AA.  
 AC 09H050;  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Probable carboxyl-terminal protease.  
 GN PA5134.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; Pubmed-10984043;  
 RA Slover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
 RA Brody L.U., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,  
 RT Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.  
 RT Nature 406:959-964(2000).  
 RL Nature 406:959-964(2000).  
 DR EMBL: AE004926; AAC08519.1; -;  
 DR MEROPS: S41.004; -;  
 DR InterPro: IPR001478; PDZ.  
 DR InterPro: IPR004447; Protease\_cterm.  
 DR InterPro: IPR003581; TSPC.  
 DR Pfam: PF00595; PDZ; 1.  
 DR SMART: SM00228; PDZ; 1.  
 DR SMART: SM00245; TSPC; 1.  
 DR TIGRFAMs: TIGR00225; prc; 1.  
 DR PROSITE: PS50106; PDZ; 1.  
 KW Protease; Complete proteome.  
 SQ SEQUENCE 436 AA; 46018 MW; D6347296673239B8 CRC64;

Query Match 70.0%; Score 35; DB 16; Length 436;  
 Best Local Similarity 63.6%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY. 1 KELAELRESTS 11  
 DB 91 EDFAELQESTS 101

Search completed: March 13, 2003, 18:51:53  
 Job time : 25.8 secs

RESULT 15  
 Q9VNM2 PRELIMINARY: PRT: 482 AA.

ID Q9VNM2  
 AC Q9VNM2  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE CG15589 protein.  
 GN CG15589.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Ephydroidea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC NCB1\_TaxID=7227;  
 OX NCB1\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 KC STRAIN-BERKELEY;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Champs M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertlera S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jajal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003600; AAF51908.1; -  
 DR FLYBase: FBgn0037409; CG15589.  
 SO SEQUENCE 482 AA; 53353 MW; E3C7E9ECC4A1226A CRC64;

Query Match 70.0%; Score 35; DB 5; Length 482;

Best Local Similarity 72.7%; Pred. No. 1.6e+02;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KELAELRESTS 11  
 DB 413 KELNELRRKPTS 423